



IE – DEPARTAMENTO DE ESTATÍSTICA  
TRABALHO DE CONCLUSÃO DE CURSO 2

**ANÁLISE DE QUALIDADE DO ENSINO NAS ESCOLAS  
PÚBLICAS DO DISTRITO FEDERAL UTILIZANDO  
PROVA BRASIL:  
UMA APLICAÇÃO DE ANÁLISE ENVOLTORIA DE  
DADOS**

Augusto Hung

**Orientador: Prof. Dr. José Angelo Belloni**

09/07/2015

Brasília  
2015



**ANÁLISE DE QUALIDADE DO ENSINO NAS ESCOLAS  
PÚBLICAS DO DISTRITO FEDERAL UTILIZANDO  
PROVA BRASIL:  
UMA APLICAÇÃO DE ANÁLISE ENVOLTORIA DE  
DADOS**

Augusto Hung

Monografia apresentada ao  
Departamento de Estatística da  
UnB como requisito parcial à  
obtenção do título de Bacharel em  
Estatística.

**Orientador: Prof. Dr. José Angelo Belloni**

Brasília  
2015



**ANÁLISE DE QUALIDADE DO ENSINO NAS  
ESCOLAS PÚBLICAS DO DISTRITO FEDERAL  
UTILIZANDO PROVA BRASIL:  
UMA APLICAÇÃO DE ANÁLISE ENVOLTORIA DE  
DADOS**

Augusto Hung

Monografia apresentada ao  
Departamento de Estatística da  
UnB como requisito parcial à  
obtenção do título de Bacharel em  
Estatística.

Aprovado em

**BANCA EXAMINADORA**

---

José Angelo Belloni  
Doutor em Engenharia de Produção na Universidade Federal de  
Santa Catarina

---

Gladston Luiz da Silva  
Doutor em Transportes na Universidade de Brasília

---

Maria Amélia Biagio  
Doutora em Engenharia Elétrica na Universidade Estadual de  
Campinas



## Resumo

As escolas públicas do Distrito Federal possuem o mesmo sistema, os alunos possuem o mesmo tipo de acesso, os professores possuem a mesma carreira. Porém os resultados da Prova Brasil são diferentes. Com isso, foram divididos em dois grandes grupos que são ensino fundamental I e II porque estão em anos diferentes que é 5º ano e 9º ano, sem esquecer que cada um desses anos faz provas diferentes. Desse modo, a cada desse grupo foi dividido em dois conglomerados que são escolas que estão localizadas em regiões mais ricas com mais infraestrutura e em regiões mais pobres com menos infraestrutura. Em seguida aplicou-se o modelo DEA (Análise Envoltória de Dados) nestes quatro grupos para descobrir as eficientes. As escolas ineficientes serão projetadas para o patamar de eficiência.

Palavra chaves: Eficiência, Conglomerado e DEA.

## Abstract

The public schools of the Federal District have the same system, students have the same access type, the teachers have the same career. But the results of the *Prova Brasil* are different. Thus, they were divided into two groups which are elementary school I and II because they are in different years is 5º and 9º, without forgetting that each of those years has different tests. Thus, each of the group was divided into two clusters which are schools that are located in richer regions with more infrastructure and in poorer areas with less infrastructure. Then it was applied the DEA (Data Envelopment Analysis) model in these four groups to discover eficiente schools. Inefficient schools will be designed to the level of efficiency.

Key word: efficiency , cluster and DEA .

# Sumário

<b>1</b>	<b>Introdução.....</b>	<b>12</b>
1.1	Objetivo Geral.....	15
1.2	Objetivos específicos.....	15
<b>2</b>	<b>Referencial teórico.....</b>	<b>16</b>
2.1	Análise por Envolvência de Dados (DEA).....	16
2.1.1	Eficiência Produtiva .....	17
2.1.2	DEA e a Eficiência Produtiva das Escolas .....	18
2.1.3	Modelo CCR.....	22
2.1.4	Modelo BCC.....	25
2.1.5	DEA e a análise de dados.....	27
2.1.6	Seleção de variáveis em DEA: .....	29
2.2	Análise de Conglomerados.....	31
<b>3</b>	<b>Metodologia.....</b>	<b>35</b>
<b>4</b>	<b>Aplicação do Modelo DEA a Escolas Públicas do DF .....</b>	<b>36</b>
4.1	Análise Exploratória dos Dados .....	36
4.1.1	Informações Disponíveis .....	36
4.1.2	Dados .....	36
4.1.3	Seleção das escolas.....	46
4.2	Modelo DEA .....	53
4.2.1	Modelo DEA 5º ano conglomerado 1.....	53
4.2.2	Modelo DEA – 5º ano Conglomerado 2 e 9º ano Conglomerados 1 e 2.....	56
<b>5</b>	<b>Aplicação dos Modelos DEA.....</b>	<b>57</b>
5.1	5º ano conglomerado 1 .....	57
5.2	5º ano, conglomerado 2 .....	58
5.3	5º ano comparando conglomerado 1 e 2. ....	60
5.4	9º ano conglomerado 1. ....	61
5.5	9º ano conglomerado 2. ....	62
5.6	Comparando as escolas eficientes do conglomerado 1 e 2. ....	63
5.7	Conclusão.....	64
<b>6</b>	<b>Conclusões e Recomendações .....</b>	<b>65</b>
<b>7</b>	<b>Referências .....</b>	<b>67</b>
<b>8</b>	<b>Anexo.....</b>	<b>68</b>



## Lista de Figuras

Figura 1: DEA – problemas dos multiplicadores .....	20
Figura 2: DEA – problemas do envelopamento .....	20
Figura 3: DEA – Medidas radiais de eficiência produtiva .....	22
Figura 4: DEA – Modelo CCR orientado para o consumo.....	23
Figura 5: DEA – Modelo CCR orientado para a produção .....	24
Figura 6: DEA – Etapas da projeção para a fronteira .....	24
Figura 7: DEA – Modelo BCC orientado para a produção .....	25
Figura 8: Fluxograma descritivo do procedimento de Norman e Stoker. ....	30
Figura 9: Nota da Prova Brasil do 5º ano da Escola Pública do Distrito Federal.....	38
Figura 10: Média ponderada da Prova Brasil do 5º ano escola pública do GDF.....	38
Figura 11: Padronização da nota do 5º ano da Prova Brasil da escola pública do DF. ....	39
Figura 12: Taxa de aprovação 5º ano escola pública do Distrito Federal. ....	40
Figura 13: Nota da Prova Brasil do 9º ano da Escola Pública do Distrito Federal.....	41
Figura 14: Média ponderada da prova brasil, 9º ano escola pública do GDF.....	41
Figura 15: Padronização da nota do 9º ano da Prova Brasil das escolas públicas do DF.....	42
Figura 16: Taxa de aprovação 9º ano escola pública do Distrito Federal. ....	43
Figura 17: Titulação dos docentes do 1º a 5º ano em 2013. ....	44
Figura 18: Título do docente 6º a 9º ano em 2013. ....	44
Figura 19: Dendograma 5º ano .....	48
Figura 20: Média padronizada das variáveis separados por conglomerado 1 e 2.....	48
Figura 21: Box Plot de taxa de aprovação segundo conglomerado 1 e 2. ....	49
Figura 22: Nota da Prova Brasil segundo conglomerados 1 e 2.....	49
Figura 23: Box Plot da Prova Brasil de Português, Matemática e média dos dois em 2011.....	50
Figura 24: Dendograma 9º ano .....	51
Figura 25: Média padronizada das variáveis separados por conglomerado 1 e 2.....	51
Figura 26: Box Plot de taxa de aprovação segundo conglomerado 1 e 2. ....	52
Figura 27: Nota da Prova Brasil segundo conglomerado 1 e 2. ....	52
Figura 28: DEA1, 5º ano, conglomerado 1 .....	53
Figura 29: DEA 2, Conglomerado 1, 5º ano.....	54
Figura 30: DEA 2, Conglomerado 1, 5º ano.....	57
Figura 31: EC 209 sul ineficiente para atingir a eficiência.....	58
Figura 32: DEA 2, Conglomerado 2, 5º ano .....	59
Figura 33: Escolas eficientes.....	60
Figura 34: DEA 2, 9º ano, conglomerado 1. ....	61
Figura 35: Escolas eficientes 9º ano, conglomerado 2.....	62
Figura 36: Escolas eficientes do conglomerado 1 e 2. ....	63

## Lista de Tabelas

Tabela 1: Número de escolas da rede pública do DF por situação urbana e rural segundo CRE.	
.....	37
Tabela 2: Número de escola no ensino fundamental por nível de ensino segundo a CRE em 2013.....	37
Tabela 3: Titulação dos professores das escolas públicas do GDF em 2013.....	43
Tabela 4: Recursos e equipamentos pedagógicos .....	45
Tabela 5: Escolas eficientes do DEA1. ....	54
Tabela 6: Escolas eficientes DEA 2, 5º ano, Conglomerado 1. ....	55
Tabela 7: Escolas eficientes DEA 2, 5º ano, Conglomerado 1. ....	57
Tabela 8: DEA 2, 5º ano, conglomerado 2.....	59
Tabela 9: 5º ano escolas eficientes .....	60
Tabela 10: escolas eficientes do 9º ano, conglomerado 1.....	61
Tabela 11: Escolas eficientes 9º ano, conglomerado 2.....	62
Tabela 12: Escolas eficientes entre conglomerado 1 e 2 .....	63
Tabela 13: Nível de Desempenho dos alunos segundo a Prova Brasil.....	68
Tabela 14: Correlação 5º ano Conglomerado 1 .....	79
Tabela 15: Escolas do 5º ano conglomerado 1 e eficientes segundo peso professor, aluno e Prova Brasil.....	80
Tabela 16: Escolas eficientes segundo escolas 5º ano conglomerado 1.....	83



## 1 Introdução.

No Brasil, os pais matriculam seus filhos na escola pública mais próxima de sua casa. A grande diferença é que as escolas de diferentes bairros são semelhantes em termos taxa de aprovação, infraestrutura e sócio econômico. Porém, elas diferem nos resultados em relação a prova padronizadas.

Sabemos como fazer uma escola de qualidade, uma escola boa. Há inclusive escolas públicas assim no Brasil, algumas centenas delas, ou talvez poucos milhares. São escolas de prestígio, de alto padrão, onde o ensino é de qualidade, os alunos estudam e aprendem e os resultados são elevados. São escolas militares, colégios de aplicação e unidades estaduais ou municipais aqui e ali que possuem as mesmas características. Mas essas escolas são poucas – uma pequena fração entre as mais de 120.000 unidades urbanas de ensino fundamental.

Nunca aprendemos a fazer aquilo que os países desenvolvidos sempre fizeram: manter um padrão. E quando o nível cai, há mecanismos para trazer a escola de volta. Resultado: embora sejam obrigados a matricular seus filhos na escola do bairro, os pais sabem que o ensino oferecido ali é semelhante ao proporcionado por unidades de outros bairros. E sabem que se seus filhos se esforçarem também obterão bons resultados.

Manter uma rede de escolas de padrão não significa que todas as unidades são idênticas, que recebem os mesmos recursos, que são 100% padronizadas. Ao contrário, para ter resultados semelhantes, as escolas precisam de recursos distintos – pois algumas têm mais problemas e desafios do que outras. Para promover a igualdade é necessário tratar desigualmente os desiguais. Escolas que caem no desempenho recebem ajuda extra; escolas com maior número de alunos com dificuldade de aprendizado recebem mais e melhores recursos, e assim por diante.

Esta pesquisa apresenta uma metodologia para avaliar a qualidade das escolas vinculadas à Secretaria de Educação do Distrito Federal sob o ponto de vista do desempenho dos seus alunos na Prova Brasil e considerando a infraestrutura disponível.

A Prova Brasil e o Sistema Nacional de Avaliação da Educação Básica (SAEB) são avaliações para diagnóstico, em larga escala, desenvolvidas pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP/MEC). Têm o objetivo de avaliar a qualidade do ensino oferecido pelo sistema educacional brasileiro a partir de testes padronizados e questionários socioeconômicos.

Nos testes aplicados na quarta e oitava séries (quinto e nono ano) do ensino fundamental, os estudantes respondem a itens (questões) de língua portuguesa, com foco em leitura, e matemática, com foco na resolução de problemas. No questionário socioeconômico, os estudantes fornecem informações sobre fatores de contexto que podem estar associados ao desempenho. Por esta razão, deve-se separar em dois grandes grupos por terem seriados e provas diferentes por serem impossível de compará-las.

Os professores e diretores das turmas e escolas avaliadas também respondem a questionários que coletam dados demográficos, perfil profissional e de condições de trabalho.

A partir das informações do Saeb e da Prova Brasil, o MEC e as secretarias estaduais e municipais de Educação podem definir ações voltadas ao aprimoramento da qualidade da

educação no país e a redução das desigualdades existentes, promovendo, por exemplo, a correção de distorções e debilidades identificadas e direcionando seus recursos técnicos e financeiros para áreas identificadas como prioritárias.

As médias de desempenho nessas avaliações também subsidiam o cálculo do Índice de Desenvolvimento da Educação Básica (IDEB), ao lado das taxas de aprovação nessas esferas.

Além disso, os dados também estão disponíveis para toda a sociedade que, a partir dos resultados, pode acompanhar as políticas implementadas pelas diferentes esferas de governo. No caso da Prova Brasil, ainda pode ser observado o desempenho específico de cada rede de ensino e do sistema como um todo das escolas públicas urbanas e rurais do país.

As informações relativas a infraestrutura física e de pessoal das escolas foram obtidas no Censo Escolar de 2013.

O Censo Escolar é um levantamento de dados estatístico-educacionais de âmbito nacional realizado todos os anos e coordenado pelo INEP. Ele é feito com a colaboração das secretarias estaduais e municipais de Educação e com a participação de todas as escolas públicas e privadas do país.

Trata-se do principal instrumento de coleta de informações da educação básica, que abrange as suas diferentes etapas e modalidades: ensino regular (educação Infantil e ensinos fundamental e médio), educação especial e educação de jovens e adultos (EJA). O Censo Escolar coleta dados sobre estabelecimentos, matrículas, funções docentes, movimento e rendimento escolar.

Essas informações são utilizadas para traçar um panorama nacional da educação básica e servem de referência para a formulação de políticas públicas e execução de programas na área da educação, incluindo os de transferência de recursos públicos como merenda e transporte escolar, distribuição de livros e uniformes, implantação de bibliotecas, instalação de energia elétrica, Dinheiro Direto na Escola e Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (FunDEB).

Além disso, os resultados obtidos no Censo Escolar sobre o rendimento (aprovação e reprovação) e movimento (abandono) escolar dos alunos do ensino Fundamental e Médio, juntamente com outras avaliações do Inep (SAEB e Prova Brasil), são utilizados para o cálculo do Índice de Desenvolvimento da Educação Básica (IDEB), indicador que serve de referência para as metas do Plano de Desenvolvimento da Educação (PDE), do Ministério da Educação.

As escolas públicas do Governo do Distrito Federal pertencem ao mesmo sistema de ensino; os professores têm a mesma carreira; o acesso dos alunos ao sistema é o mesmo e os recursos provem da mesma fonte. As escolas, porém, têm alguma autonomia na alocação de parte destes recursos e na execução do processo pedagógico. Porém, observou-se diferenças significativas no desempenho dos alunos avaliados na Prova Brasil.

Para mensuração de desempenho das escolas utilizou-se a Análise Envoltória de Dados, técnica que avalia a eficiência com que a escola transforma os recursos disponíveis (professores, titulação do professor, alunos matriculados, equipamentos) em resultados (desempenho da Prova Brasil, taxa de aprovação).

A técnica DEA considera múltiplos recursos e múltiplos resultados sem exigir que se conheça a priori um conjunto de taxas de substituição (pesos) entre as variáveis nem a

especificação da forma funcional das relações entre os recursos e os resultados. Estas características são particularmente vantajosas na avaliação de instituições cujos fatores de produção não estão diretamente submetidos a valores de mercado ou outras medidas cardinais de importância, como, por exemplo, escolas.

Apesar de pertencerem ao mesmo sistema de ensino, as escolas do Distrito Federal diferenciam-se com relação a infraestrutura disponível além de localizarem-se em regiões administrativas com situações socioeconômicas bastante distintas.

Para garantir a comparabilidade entre as escolas, condição necessária para a aplicação da DEA, classificou-se as escolas em grupos com características semelhantes, utilizando Análise de Conglomerados (*Cluster Analysis*).

Em cada segmento analisado (Ensino Fundamental 1 e Ensino Fundamental 2) identificou-se dois conglomerados com características distintas, uns formados por escolas com boas condições de infraestrutura localizadas em regiões de rendas mais altas e outra com piores condições de infraestrutura em regiões de renda mais baixa.

Aplicou-se DEA a cada um dos conglomerados em cada segmento, classificando, assim, as escolas em eficientes e ineficientes e identificando possibilidade de melhoria para as ineficientes.

Num segundo momento comparou-se as escolas de ambos os conglomerados em cada segmento, observando-se a permanência de escolas menos equipadas como eficientes. Essas escolas constituem referências de bom desempenho dos alunos mesmo em condições desfavoráveis, e devem ser objeto de investigações futuras.

A projeção das escolas ineficientes nas fronteiras de eficiência permitiu, não só mensurar as ineficiências, mas, também, identificar ações e estratégias para aumentar a produtividade.

### **1.1 Objetivo Geral**

O objetivo deste trabalho é avaliar a qualidade da educação das escolas públicas do Distrito Federal, fazendo a análise de recursos como financeiro, patrimonial, humano e fazer uma dualidade entre a Prova Brasil e a quantidade de alunos aprovados no ensino fundamental.

### **1.2 Objetivos específicos**

Realizar uma análise descritiva do sistema público de ensino do Distrito Federal.

Apresentar uma metodologia para avaliar o desempenho das escolas públicas do DF sob o ponto de vista da eficiência produtiva, considerando indicadores de qualidade do ensino.

Identificar estratégias e ações que conduzam as Escolas ineficientes ao patamar da eficiência respeitando seus projetos pedagógicos e especificidades ambientais

## 2 Referencial teórico

### 2.1 Análise por Envoltória de Dados (DEA)

A tese de Doutorado de Edward Rhodes, orientado por W.W. Cooper, deu origem à Análise por Envoltória de Dados, na qual o problema abordado foi o desenvolvimento de um método para comparar a eficiência de escolas públicas – DMU's levando em consideração os outputs (Resultados) segundo Charnes et al. (1994). Esta tese tinha como objetivo desenvolver um modelo para estimar a eficiência técnica, na qual não se utilizava pesos para as variáveis, tanto para os inputs (Recursos) quanto para os outputs (Resultados) e sem converter todas as variáveis em valores econômicos comparáveis (Modelo CCR).

O que diferencia o método DEA dos demais métodos baseados em avaliação puramente econômica é que estes necessitam converter todos os RECURSOS e RESULTADOS em unidades monetárias. O DEA irá tratar simultaneamente múltiplos recursos e múltiplos resultados sem exigir, a priori, nenhuma estrutura de pesos entre eles.

Além disso, DEA considera outliers, desvios em relação ao comportamento médio, como possíveis benchmarks a serem tomados como referência pelas demais DMU's, avalia cada observação individual com o objetivo de determinar uma fronteira linear por partes, que compreende o conjunto de DMU's Pareto-Eficiente.

DEA é uma técnica utilizada por profissionais de várias áreas do conhecimento com os enfoques de interesse (Seiford, 1999; apud LINS; MEZA, 2000).

Os econométricos consideram DEA como uma técnica para estimação de uma função de produção empírica. Os matemáticos a consideram uma metodologia para determinar soluções não dominadas em um problema com múltiplos critérios. E para os engenheiros industriais, DEA constitui uma ferramenta para o aumento e melhoria da produtividade.



### 2.1.1 Eficiência Produtiva

Sob aspectos econômicos e de produtividade material a eficiência na produção se relaciona com a capacidade que uma organização tem de transformar os recursos no máximo de resultados.

A eficiência produtiva é a capacidade de se evitar desperdícios afim de produzir a maior quantidade de resultados que os recursos utilizados permitem e utilizar o mínimo possível de recursos necessários à determinada produção.

Um modelo de avaliação de eficiência conhecido como modelo CCR (Retorno Constantes a Escalas – CRS) foi proposto por Charnes, Cooper e Rhodes (1978), dando origem a um complexo de modelos e técnicas de construção de fronteiras de produção e medidas de eficiência relativa, conhecida como Análise por Envoltória de Dados (DEA).

O modelo conhecido como BCC (Retorno Variáveis a Escalas– VRS) desenvolvido por Banker, Charnes e Cooper em 1984, é uma medida de eficiência técnica, isolando o componente da eficiência produtiva total que é devido a afastamentos da escala mais produtiva (eficiência de escala) do componente de eficiência técnica propriamente dita (eficiência técnica). (BELLONI, 2000)

O DEA é um método largamente usado para estimar as eficiências de unidades organizacionais que possuem as mesmas características e que usam um mesmo conjunto de recursos para produzir um mesmo conjunto de resultados, por meio de processos tecnológicos semelhantes.

O modelo DEA considera valores observados de recursos utilizados e resultados alcançados para construir um espaço de possibilidades de produção, delimitado por uma fronteira linear por partes definida a partir dos melhores desempenhos. As unidades que se posicionam sobre essa fronteira são ditas eficientes, enquanto a ineficiência das demais é mensurada a partir da medida da distância à fronteira. Essa distância determina a possibilidade de aumento na produção.

DEA é adequado quando não é possível ou apropriado unir recursos e resultados em uma única unidade.

### 2.1.2 DEA e a Eficiência Produtiva das Escolas

A eficiência produtiva mensurada pela técnica DEA será conceituada formalmente. Considere uma escola como um sistema de produção múltipla, que transforma  $N$  itens de recursos, representados por um vetor de quantidades  $x = (x_1, x_2, \dots, x_N) \in R_+^N$ , em  $M$  itens de resultados cujas quantidades estão representadas em um vetor  $y = (y_1, y_2, \dots, y_M) \in R_+^M$ , determinando um plano de produção descrito pelo vetor  $(x, y) \in R_+^{N+M}$ .

Supondo que foram observados  $K$  planos de produção  $(x_k, y_k), k = 1, 2, \dots, K$ , realizados por  $K$  escolas similares, que foram denotados por *DMU's* (Decision Making Unity). O  $x_{ki}$  é a quantidade do recurso  $i$  ( $i = 1, 2, \dots, N$ ) utilizada e  $y_{kj}$  é a quantidade do resultado  $j$  ( $j = 1, 2, \dots, M$ ) produzida pela  $DMU_k$ . Neste trabalho, a escola avaliada será simbolizada por  $DMU_0$ , representada nos modelos DEA pelo plano de produção  $(x_0, y_0)$ .

A medida de desempenho para cada  $DMU_0$  é calculada pela Análise por Envoltória de

$$E_0 = \frac{\sum_{j=1}^M q_j y_{0j}}{\sum_{i=1}^N p_i x_{0i}},$$

Dados que tem a seguinte expressão:

onde  $p = (p_1, p_2, \dots, p_N)$  e  $q = (q_1, q_2, \dots, q_M)$  são os vetores de pesos utilizados para a agregação dos recursos e dos resultados da  $DMU_0$ , respectivamente. O desempenho resultante,  $E_0$ , é uma medida da produtividade da  $DMU_0$ , por configurar-se uma razão entre a produção agregada e o consumo agregado.

A utilização de um conjunto de pesos para cada  $DMU$  resulta em uma medida de desempenho específica para cada escola sob avaliação.

Os pesos para cada  $DMU_0$  são determinados maximizando-se o valor da produtividade  $E_0$ , sujeito à restrição de que a produtividade de nenhuma  $DMU$ , calculada com os pesos da  $DMU_0$ , possa exceder um valor constante pré-fixado (usualmente tomado igual a 1). O seguinte problema de programação matemática pode ser usado para calcular tais pesos:

$$E_0 = \max \frac{\sum_{j=1}^M q_j y_{0j}}{\sum_{i=1}^N p_i x_{0i}}$$

$$s/a \quad E_k = \frac{\sum_{j=1}^M q_j y_{kj}}{\sum_{i=1}^N p_i x_{ki}} \leq 1 \quad \forall k = 1, 2, \dots, K \quad (1)$$

$$p_i \geq 0, \quad \forall i = 1, 2, \dots, N; \quad q_j \geq 0, \quad \forall j = 1, 2, \dots, M$$

A medida  $E_0$ , produtividade da  $DMU_0$ , pode ser dividida pela produtividade máxima, constituindo-se, assim, numa medida da eficiência relativa da  $DMU_0$ , já que a produtividade máxima observada será sempre igual a 1.

Na solução desse problema de otimização, a eficiência da  $DMU_0$  é maximizada sob a condição que a eficiência de cada uma das unidades não excede o valor 1. Assim, uma  $DMU_k$  será considerada eficiente, sob o ponto de vista da  $DMU_0$ , quando sua medida de eficiência relativa  $E_k$  (calculada com os pesos da  $DMU_0$ ) for igual a 1, e ineficiente quando esta medida for menor que 1.

Na equação (1), o modelo definido caracteriza-se por um problema de programação fracional que pode ser reduzido a dois problemas de programação linear, empregando o procedimento de transformação de problemas fracionais de Charnes e Cooper (1962, apud Belloni, 2000). Essa transformação se faz alterando a função objetivo em (1) e resulta em dois problemas de programação linear: Um que mantém constante o agregado de recursos e busca maximizar o agregado de produção (2); e, outro, que minimiza o agregado de recursos, mantendo constante o agregado de produção (3). Esses problemas estão explicitados na Figura 1.

Os problemas (2) e (3) são chamados problemas dos multiplicadores, por expressarem as taxas de substituição entre recursos e entre resultados (os pesos) – que definem a faceta da fronteira de eficiência na qual é projetada a  $DMU_0$ . O conceito de desempenho que está sendo mensurado em cada um deles torna-se mais claro quando se analisam problemas equivalentes a (2) e (3). Os problemas (4) e (5) da Figura 2 correspondem aos problemas duais dos problemas (2) e (3), respectivamente, e são conhecidos como problemas do envelopamento.

**Figura 1: DEA – problemas dos multiplicadores**

DEA orientado para o consumo (problema dos multiplicadores) ( 2 )	DEA orientado para a produção (problema dos multiplicadores) ( 3 )
$\max_{p_i, q_j} \sum_{j=1}^M y_{0j} q_j$ <p>s/a</p> $\sum_{i=1}^N x_{0i} p_i = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq 0, \quad \forall i = 1, 2, \dots, N$ $q_j \geq 0 \quad \forall j = 1, 2, \dots, M$	$\min_{p_i, q_j} \sum_{i=1}^N x_{0i} p_i$ <p>s/a</p> $\sum_{j=1}^M y_{0j} q_j = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq 0, \quad \forall i = 1, 2, \dots, N$ $q_j \geq 0 \quad \forall j = 1, 2, \dots, M$

**Figura 2: DEA – problemas do envelopamento**

DEA orientado para o consumo (problema do envelopamento) ( 4 )	DEA orientado para a produção (problema do envelopamento) ( 5 )
$\min \theta$ <p>s/a</p> $\sum_{k=1}^K z_k y_{kj} \geq y_{0j}, \quad \forall j = 1, 2, \dots, M$ $x_{0i} \theta - \sum_{k=1}^K z_k x_{ki} \geq 0 \quad \forall i = 1, 2, \dots, N$ $\theta \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$	$\max \lambda$ <p>s/a</p> $y_{0j} \lambda - \sum_{k=1}^K z_k y_{kj} \leq 0, \quad \forall j = 1, 2, \dots, M$ $\sum_{k=1}^K z_k x_{ki} \leq x_{0i}, \quad \forall i = 1, 2, \dots, N$ $\lambda \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$

As regiões de viabilidade dos problemas (4) e (5) caracterizam, respectivamente, o conjunto de necessidades de consumo associado ao vetor de produção da  $DMU_0(L(y_0))$  e o

conjunto de possibilidades de produção associado ao seu vetor de consumo ( $P(x_0)$ ). Os escalares  $Z_k$  são os coeficientes dos planos de produção  $(x_k, y_k)$ , nas combinações lineares que definem a tecnologia de produção.

O valor ótimo para  $\theta$  no problema (4), que notaremos  $\theta^*$ , representa a contração equiproporcional máxima possível no vetor de recursos da  $DMU_0$ , mantendo-se constante o vetor de resultados observados. Se nenhuma contração equiproporcional for possível, então  $\theta^*=1$  e a  $DMU_0$  é dita eficiente na isoquanta do conjunto de necessidades de consumo  $L(y_0)$ . Se  $\theta^*<1$ , então a  $DMU_0$  é ineficiente e  $\theta^*$  é a medida de sua ineficiência, já que os recursos utilizados poderiam ser reduzidos equiproporcionalmente de  $x_0$  para  $\theta * x_0$  sem redução nos resultados produzidos.

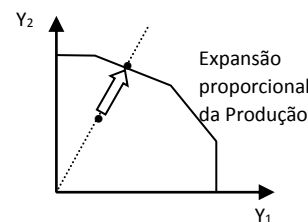
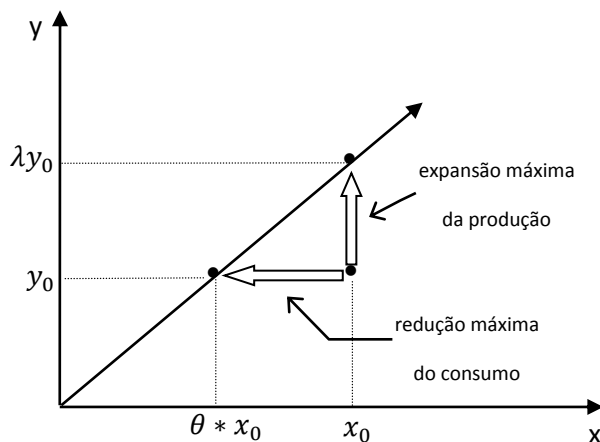
De maneira equivalente,  $\lambda^*$ , o valor ótimo do problema (5), é a expansão máxima possível no vetor de resultados da  $DMU_0$ , mantendo-se constante o vetor de recursos utilizados. Se  $\lambda^*=1$ , então nenhuma expansão é possível e a  $DMU_0$  é dita eficiente na isoquanta do conjunto de possibilidades de produção  $P(x_0)$ . Se  $\lambda^*>1$ , então a  $DMU_0$  é dita ineficiente e seus resultados podem ser expandidos de  $y_0$  para  $\lambda * y_0$  sem acréscimo de recursos. O valor  $1/\lambda^*$  define a medida de eficiência fraca da  $DMU_0$  orientada para o aumento da produção de resultados.

As medidas radiais (proporcionais) de eficiência estão ilustradas na Figura 3, que mostra as orientações para a maximização da produção e para a minimização do consumo e as projeções radiais (equiproporcionais) para as fronteiras.

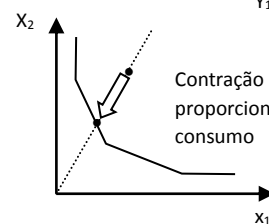
As medidas definidas nos problemas (4) e (5) se caracterizam por uma projeção radial (equiproporcional) sobre a fronteira, mantendo, portanto, as proporções entre recursos e entre resultados observadas na  $DMU_0$ .

As medidas radiais (equiproporcionais) têm a vantagem de ser independentes de unidades de medidas e, portanto, independentes de relações de preços. Por outro lado, a contração equiproporcional de todos os recursos, ou expansão equiproporcional de todos os resultados, sugere eficiência mesmo quando permanecem excessos no consumo de algum recurso ou folga (falta) na produção de algum resultado.

**Figura 3: DEA – Medidas radiais de eficiência produtiva**



**Figura 3.1**



**Figura 3.2**

### 2.1.3 Modelo CCR

É sempre possível trabalhar, separadamente, com a medida de eficiência radial e os possíveis excessos no consumo de recursos e folgas na produção. Charnes, Cooper e Rhodes (1978) propõem um modelo de avaliação da eficiência que calcula, simultaneamente, as folgas e os excessos existentes, mesmo mantendo o princípio da projeção radial sobre a fronteira (Modelo CCR). A Figura 4 (orientado para o consumo de recursos) e Figura 5 (orientado para a produção de resultados) representam os problemas de programação linear associados ao modelo CCR.

Apesar do cálculo em paralelo da expansão proporcional e dos excessos e folgas, o procedimento de otimização proposto pode ser interpretado em duas etapas consecuentes:

- i. um deslocamento radial na direção da fronteira e,
- ii. um movimento não-radial para a fronteira de eficiência.

Dessa forma, a proporção de uma *DMU* ineficiente sobre a fronteira é feita em duas etapas conforme a Figura 6.

As expressões  $\sum_{k=1}^K z_k y_{kj}$  e  $\sum_{k=1}^K z_k x_{ki}$  nas restrições dos problemas do envelopamento representam combinações lineares dos resultados e dos recursos das *KDMU's* tomadas como referências e o escalar  $z_k$  é o coeficiente da *DMU<sub>k</sub>* nessas combinações. A falta de restrições sobre os valores dos  $z_k$  pressupõe a hipótese de retornos constantes à escala de produção, uma vez que quaisquer expansões ou contrações dos planos de produção observados sejam viáveis. As desigualdades presentes nas duas primeiras restrições desses problemas caracterizam o livre descarte de recursos e resultados.

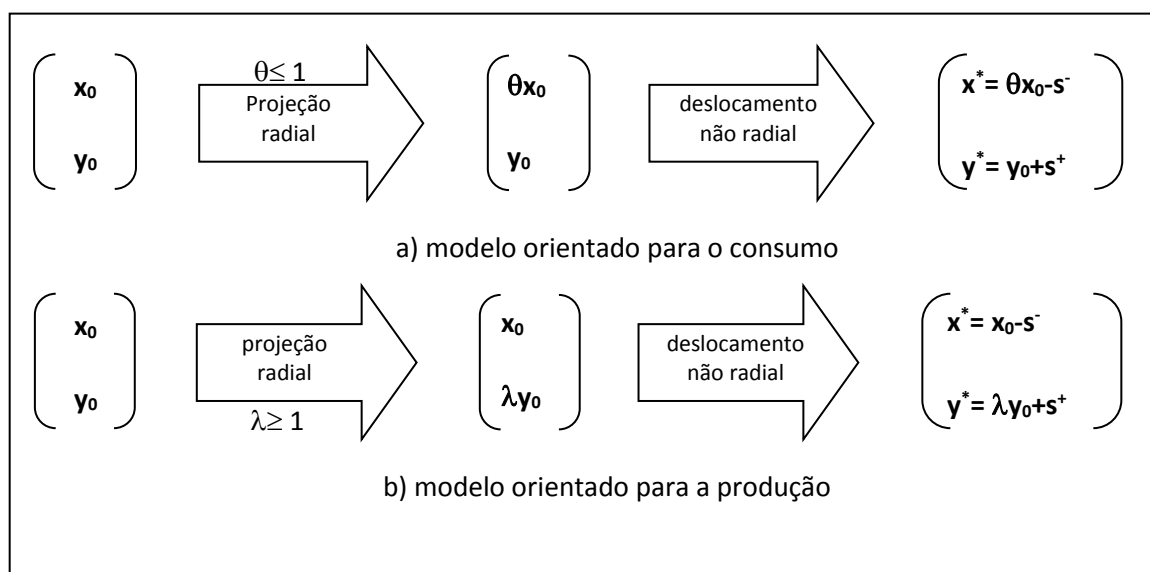
**Figura 4: DEA – Modelo CCR orientado para o consumo.**

Modelo CCR orientado para o consumo retornos constantes à escala - descarte forte de recursos e resultados	
(problema dos multiplicadores)	(problema do envelopamento)
$\max_{p_i, q_j} \sum_{j=1}^M y_{0j} q_j$ <p style="text-align: center;">s/a</p> $\sum_{i=1}^N x_{0i} p_i = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq \varepsilon, \quad \forall i = 1, 2, \dots, N$ $q_j \geq \varepsilon \quad \forall j = 1, 2, \dots, M$ $\varepsilon > 0$	$\min \quad \theta - \varepsilon (\sum_{j=1}^M s_j^+ + \sum_{i=1}^N s_i^-)$ <p style="text-align: center;">s/a</p> $\sum_{k=1}^K z_k y_{kj} - s_j^+ = y_{0j}, \quad \forall j = 1, 2, \dots, M$ $x_{0i} \theta - \sum_{k=1}^K z_k x_{ki} - s_i^- = 0, \quad \forall i = 1, 2, \dots, N$ $\theta \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$ $s_j^+ \geq 0, \quad \forall j = 1, 2, \dots, M; \quad s_i^- \geq 0, \quad \forall i = 1, 2, \dots, N$

**Figura 5: DEA – Modelo CCR orientado para a produção**

Modelo CCR orientado para a produção retornos constantes à escala - descarte forte de recursos e resultados	
(problema dos multiplicadores)	(problema do envelopamento)
$\min_{p_i, q_j} \sum_{i=1}^N x_{0i} p_i$ <p style="text-align: center;">s/a</p> $\sum_{j=1}^M y_{0j} q_j = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq \varepsilon, \quad \forall i = 1, 2, \dots, N$ $q_j \geq \varepsilon \quad \forall j = 1, 2, \dots, M$ $\varepsilon > 0$	$\max \lambda + \varepsilon (\sum_{j=1}^M s_j^+ + \sum_{i=1}^N s_i^-)$ <p style="text-align: center;">s/a</p> $y_{0j} \lambda - \sum_{k=1}^K z_k y_{kj} + s_j^+ = 0, \quad \forall j = 1, 2, \dots, M$ $\sum_{k=1}^K z_k x_{ki} + s_i^- = x_{0i}, \quad \forall i = 1, 2, \dots, N$ $\lambda \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$ $s_j^+ \geq 0, \quad \forall j = 1, 2, \dots, M; \quad s_i^- \geq 0, \quad \forall i = 1, 2, \dots, N$

**Figura 6: DEA – Etapas da projeção para a fronteira**





### 2.1.4 Modelo BCC

Um modelo DEA foi desenvolvido por Banker, Charnes e Cooper (1984), conhecido como Modelo BCC, que pressupõe tecnologias que exibam retornos variáveis à escala de produção e descarte forte de recursos e resultados. Esse modelo admite que a produtividade máxima varie em função da escala de produção, ao possibilitar que a tecnologia exiba propriedades de retornos à escala diferentes ao longo de sua fronteira.

A medida de eficiência resultante permite que se isole da ineficiência produtiva total o componente associado à escala de produção, permitindo identificar a ineficiência técnica. O modelo possibilita a utilização de unidades de referência de portes distintos viabilizando assim o uso de todas as escolas como unidades de referência, independentemente do tamanho, pois estas estão livres das dificuldades que surgem ao considerar a escala de produção.

Os problemas de programação linear associados ao Modelo BCC estão representados na Figura 7, na qual estão assinaladas as diferenças em relação ao Modelo CCR.

**Figura 7: DEA – Modelo BCC orientado para a produção**

Modelo BCC orientado para a produção retornos variáveis à escala - descarte forte de recursos e resultados	
$\min_{p_i, q_j} \sum_{i=1}^N x_{0i} p_i + v_0$ <p style="text-align: right;">Retornos variáveis</p> <p style="text-align: center;">s/a</p> $\sum_{j=1}^M y_{0j} q_j = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j + v_0 \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq \varepsilon, \quad \forall i = 1, 2, \dots, N$ $q_j \geq \varepsilon \quad \forall j = 1, 2, \dots, M$ $\varepsilon > 0$ <p style="text-align: center;">(problema dos multiplicadores)</p>	$\max \lambda + \varepsilon (\sum_{j=1}^M s_j^+ + \sum_{i=1}^N s_i^-)$ <p style="text-align: center;">s/a</p> $y_{0j} \lambda - \sum_{k=1}^K z_k y_{kj} + s_j^+ = 0, \quad \forall j = 1, 2, \dots, M$ $\sum_{k=1}^K z_k x_{ki} + s_i^- = x_{0i}, \quad \forall i = 1, 2, \dots, N$ $\sum_{k=1}^K z_k = 1$ <p style="text-align: right;">Retornos variáveis</p> $\lambda \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$ $s_j^+ \geq 0, \quad \forall j = 1, 2, \dots, M; \quad s_i^- \geq 0, \quad \forall i = 1, 2, \dots, N$ <p style="text-align: center;">(problema do envelopamento)</p>

O problema do envelopamento do Modelo BCC se diferencia do seu equivalente no Modelo CCR pela restrição  $\sum_{k=1}^K z_k = 1$ , que restringe as combinações lineares dos planos observados a combinações convexas desses planos. A tecnologia restringe-se a contrações dos planos de produção observados, impedindo expansões e, assim, caracterizando a hipótese de retornos variáveis à escala.

O hiperplano suporte ao conjunto tecnologia definido por  $(p^*, q^*)$ , no problema dos multiplicadores, tem um termo independente de variável ( $v_0$ ) que exerce o papel de intercepto, possibilitando a existência na fronteira de facetas definidas por hiperplanos que não passam na origem, o que caracteriza, de forma equivalente ao problema do envelopamento, retornos variáveis à escala.

### 2.1.5 DEA e a análise de dados

Considerando qualquer um dos modelos citados de DEA, cada escola será avaliada sob a visão que mais lhe convém (os pesos relativos que maximizam sua produtividade) e só será considerada ineficiente se outra(s) escola(s) obtiver(em), com a estrutura de pesos escolhida, uma produtividade maior. Desta forma, para cada escola avaliada, se estabelece uma ordenação de todas elas, segundo as eficiências calculadas com os pesos relativos adotados (resultantes do processo de otimização).

O modelo DEA empregado para cada escola ineficiente identifica um conjunto de escolas eficientes que formam um grupo de referência (*benchmarks*) para a análise do desempenho da escola em avaliação. Essas escolas de referência determinam uma faceta no conjunto de possibilidades de produção na qual será projetado (Figura 6).

O vetor de resultados  $y_0$  é expandido radialmente para  $\lambda * y_0$ , mantido o vetor de recursos  $x_0$ , determinando uma meta (meta proporcional) para a produção constitui a primeira etapa. Um novo conjunto de relações parciais de produtividade é determinado e este explicita as relações entre recursos e resultados que podem ser alcançadas sem alteração nas quantidades de recursos consumidos.

Numa segunda etapa, serão tratadas as possíveis folgas na produção e excessos no consumo observados após a expansão proporcional. A eliminação das folgas e excessos é proposta a partir de um conjunto de pesos (taxas de substituição) fornecido pelo problema dos multiplicadores, obtendo-se uma meta global para cada escola ineficiente.

Conhecendo a medida de eficiência, a análise dos resultados da aplicação do modelo DEA se concentra no estudo da faceta de eficiência associada a cada escola avaliada. As características da faceta de eficiência permitem estudar o sentido da medida obtida e a viabilidade das metas projetadas para cada escola.

A medida de eficiência relativa de cada escola calculada pelos modelos DEA corresponde à distância entre o plano de operação realizado e a fronteira eficiente. A fronteira é composta de seções lineares – chamadas facetas de eficiência – determinadas pelas combinações convexas de subconjuntos de escolas eficientes. Cada escola corresponde a uma faceta de eficiência em relação à qual sua medida de eficiência é calculada e na qual sua meta global é projetada.

As facetas de eficiência são caracterizadas por um conjunto de vértices e um conjunto de taxas de substituição entre variáveis. Os vértices definem os limites de cada faceta,

enquanto as taxas de substituição refletem relações de valor entre as variáveis e descrevem as condições de equilíbrio (otimalidade) que garantem a condição de eficiência.

Os vértices da faceta de eficiência são as escolas eficientes de referência para a escola avaliada. A meta projetada para cada escola ineficiente é uma combinação convexa dos planos de operação de suas escolas de referência. As referências caracterizam-se como planos de operação similares a cada escola a ser avaliada.

As taxas de substituição estão expressas nas variáveis duais e representam produtividades marginais que consistem em ponderadores utilizados na agregação dos recursos e dos resultados. Essas taxas definem as funções de produção que determinam a fronteira de eficiência. Para cada escola avaliada, as taxas de substituição descrevem as relações de produção associadas à sua faceta na fronteira de possibilidades de produção, definindo a função de produção naquela seção linear da fronteira.

Na avaliação do desempenho de uma escola, na qual não existem preços relativos entre recursos e entre resultados, as taxas de substituição devem ser interpretadas como indicadores da importância relativa dos recursos e dos resultados na caracterização do plano de operação eficiente adotado como meta global.

Para avaliar a importância de cada escola eficiente na determinação das facetas é necessário identificar o número de vezes que cada uma delas aparece como referência para os ineficientes.

### 2.1.6 Seleção de variáveis em DEA:

#### Procedimento de seleção de Norman e Stoker

Norman e Stoker (1991) propuseram um procedimento de seleção de variáveis para a construção de funções de desempenho que utiliza uma combinação de análise de correlações lineares simples e DEA.

A aplicação da técnica DEA pressupõe uma classificação prévia das variáveis em recursos ou resultados para a construção da função de desempenho. Uma análise estatística dos dados deve identificar as relações existentes entre as variáveis e entre conjuntos de variáveis, condição necessária para classificá-las como descritoras dos recursos ou dos resultados do processo produtivo. As medidas de associação estatística entre recursos e resultados devem refletir relações causais adequadas, identificando, em particular, que recursos são responsáveis por cada resultado e que resultados são esperados de cada recurso.

Devem, também, serem analisadas as associações entre recursos e entre resultados para identificar se existe complementaridade ou competição entre eles. Assim, não deve haver independência estatística entre duas variáveis descritoras dos resultados ou dos recursos, isto é, a associação entre duas variáveis representativas dos recursos ou ambas representativas dos resultados deve ser positiva (caracterizando atividades complementares) ou negativa (atividades concorrentes ou competitivas), mas não deve ser (muito próxima de) zero.

Conhecida a classificação das variáveis em recursos ou resultados, utiliza-se o procedimento de seleção de variáveis proposto por Norman e Stoker (1991) para identificar aquelas que mais influenciam o desempenho de uma universidade e construir um indicador único da eficiência produtiva que considera esses fatores conjuntamente. Esse procedimento identifica os recursos e os resultados mais relevantes à mensuração do desempenho adotado através da construção de uma sequência de funções de desempenho.

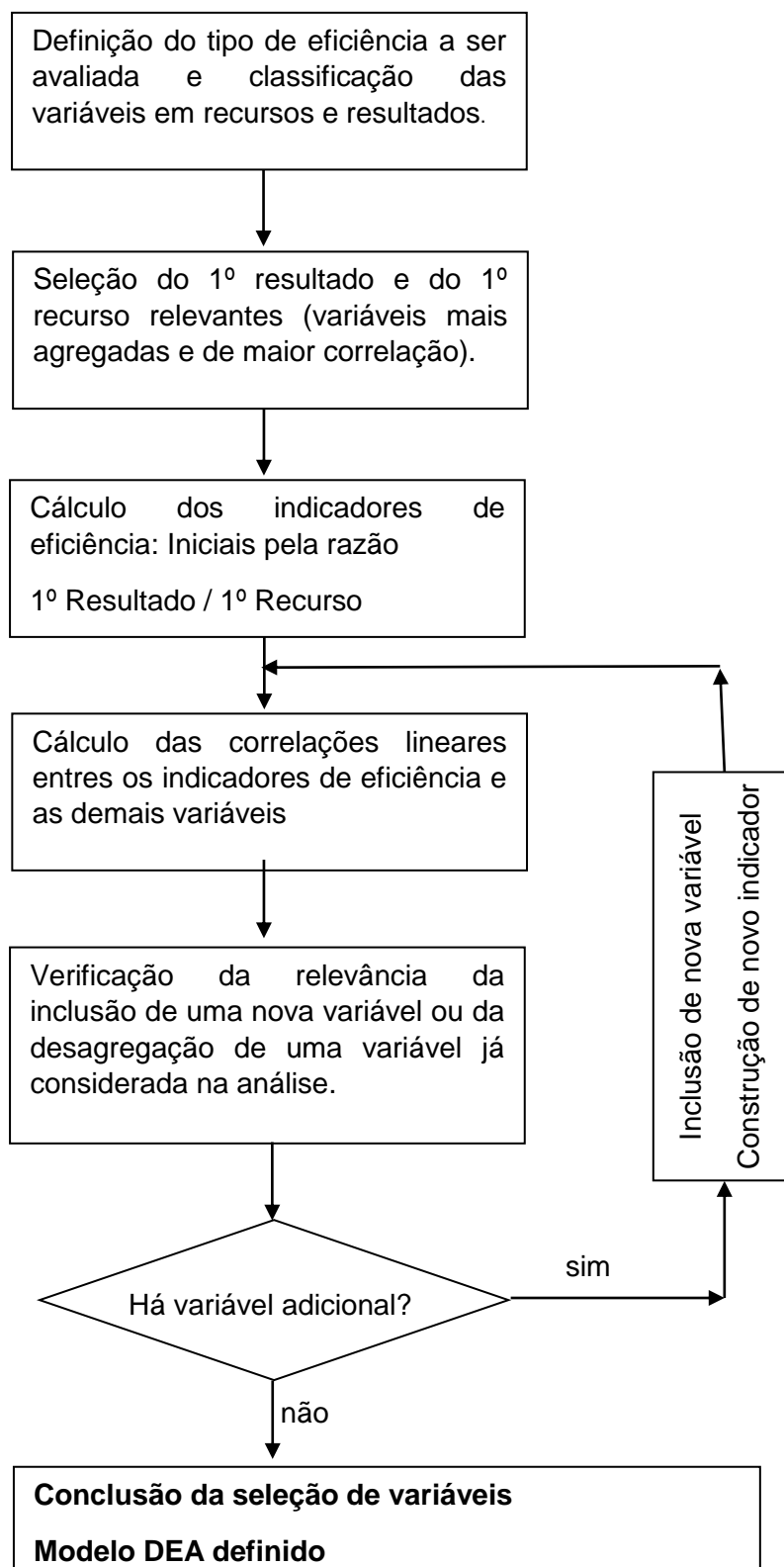
A função de desempenho da primeira etapa do procedimento de Norman e Stoker corresponde a um único resultado e um único recurso. Em cada uma das etapas seguintes são observadas as associações entre cada variável e o indicador da eficiência da etapa anterior para decidir se existem novas variáveis a considerar na função de desempenho. Essas novas variáveis podem representar um novo recurso ou um novo resultado, bem como decomposições de variáveis já presentes no modelo.

Se alguma variável que influencia a eficiência estiver ausente da função de desempenho, então o indicador da eficiência calculado deverá estar viesado com relação ao fator educacional descrito por esta variável. Deve-se esperar, nesse caso, associação estatística significativa entre o indicador da eficiência e essa variável, que é indicativa da possibilidade de entrada dessa variável na função de desempenho. A recíproca não é, necessariamente, verdadeira. Associações estatísticas fortes não permitem concluir que uma dada variável influencia o desempenho. É necessário determinar relações causais lógicas entre cada fator e a eficiência produtiva.

A relevância da decomposição de variáveis que estão presentes de forma agregada na função de desempenho é verificada através da análise das correlações entre as suas parcelas e

o indicador da eficiência. Associações inversas entre parcelas e a eficiência são sinalizadoras da necessidade de decomposição da variável agregada.

**Figura 8: Fluxograma descritivo do procedimento de Norman e Stoker.**



## 2.2 Análise de Conglomerados

A análise de conglomerados consiste em um processo de partição de uma população heterogênea em vários subgrupos mais homogêneos, seu objetivo é buscar entender a estrutura dos dados aglomerando observações similares em grupos. No agrupamento, não há classes pré-definidas, os elementos são agrupados de acordo com a similaridade entre eles. Os grupos são determinados de forma a obter-se homogeneidade dentro dos grupos e heterogeneidade entre eles.

A sua aplicação é, portanto, bastante imediata e pode conferir um bom poder explanatório à pesquisa, supondo que decisões ponderadas sejam tomadas sobre formas alternativas para mensuração de similaridades, critérios para formação de grupos e número ideal de grupos a serem formados.

Uma análise de conglomerados mais criteriosa deve apresentar métodos com as seguintes características:

- Ser capaz de lidar com dados com alta dimensionalidade;
- Ser “escalável” com o número de dimensões e com a quantidade de elementos a serem agrupados;
- Habilidade para lidar com diferentes tipos de dados;
- Capacidade de definir agrupamentos de diferentes tamanhos e formas;
- Ser robusto à presença de ruído;
- Apresentar resultado consistente independente da ordem em que os dados são apresentados.

Existem várias técnicas de análise de conglomerados, as quais permitem diferentes medidas de similaridade e métodos de formação de grupos. As medidas de similaridades entre os elementos a serem agrupados normalmente são expressas como uma função distância ou métrica.

Seja  $M$  um conjunto um conjunto de dados, aplicando-se uma métrica em  $M$  tem-se uma função  $d: M \times M \rightarrow \mathbb{R}$ , tal que para qualquer observação  $A, B, C \in M$  a métrica deve satisfazer:

1.  $d(A, B) > 0$ , para todo  $A \neq B$
2.  $d(A, A) = 0$
3.  $d(A, B) = d(B, A)$
4.  $d(A, B) \leq d(A, C) + d(C, B)$

Serão apresentadas, a seguir, algumas medidas que calculam as similaridades, que são as distâncias entre os elementos. Essa distância é normalmente representada na forma de matriz, ou seja, em uma matriz de similaridade. A matriz de similaridade é simétrica.

A distância euclidiana é a distância geométrica no espaço multidimensional. A distância euclidiana entre dois elementos  $X = [X_1, X_2, \dots, X_p]$  e  $Y = [Y_1, Y_2, \dots, Y_p]$  é dada por:

$$d_{xy} = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + \dots + (X_p - Y_p)^2} = \sqrt{\sum_i^p (X_i - Y_i)^2}$$

A distância Manhattan é definida pela expressão:

$$d_{xy} = \sqrt{|X_1 - Y_1| + |X_2 - Y_2| + \dots + |X_p - Y_p|} = \sqrt{\sum_i^p |X_i - Y_i|}$$

Em alguns casos os resultados apresentados pela distância de Manhattan são próximos aos da distância Euclidiana. No caso da distância de Manhattan como as diferenças não são elevadas ao quadrado o efeito das dimensões elevadas de algum elementos é maximizado.

A distância de Chebychev é apropriada no caso em que se deseja definir dois elementos como diferentes, se apenas uma das dimensões é diferente. Ela é definida por:

$$d_{xy} = \text{máximo}(|X_1 - Y_1|, |X_2 - Y_2|, \dots, |X_p - Y_p|)$$

Existem várias técnicas que são utilizadas para formar os grupos, ou seja, são técnicas que auxiliam na formação dos conglomerados.



As Técnicas Hierárquicas Aglomerativas consideram cada indivíduo (observação) como um conglomerado de tamanho 1, ou seja, inicialmente têm-se  $n$  conglomerados. Os pares de conglomerados mais similares são agrupados e passam a constituir um único conglomerado.

Os métodos aglomerativos são caracterizados de acordo com o critério utilizado para medir a distância entre os grupos. Eis alguns deles:

1) Métodos de ligação:

- *Single linkage* (ou ligação por vizinho mais próximo);
- *Complete linkage* (ou ligação por vizinho mais distante);
- *Average linkage* (ou ligação por média);
- *Median linkage* (ou ligação por mediana);

2) Métodos de centroide (ou ligação por centróide);

3) Métodos de soma de erros quadráticos ou variância (método de *Ward*).

Os grupos, nos métodos hierárquicos, são geralmente representados por um diagrama bi-dimensional chamado de dendograma ou diagrama de árvore. Neste diagrama, cada ramo representa um elemento, enquanto a raiz representa o agrupamento de todos os elementos. Através do dendograma e do conhecimento prévio sobre a estrutura dos dados, deve-se determinar uma distância de corte para definir quais serão os grupos formados.

Essa decisão é subjetiva, e deve ser feita de acordo o objetivo da análise e o número de grupos desejados. Se eles apresentam resultados semelhantes, significa que o grupo está bem definido no espaço, ou seja, o grupo é real. Mas se ocorre o contrário, os grupos provavelmente não existem (ROMESBURG, 1984).

As Técnicas Hierárquicas Divisivas trabalham na direção oposta dos métodos aglomerativos, ou seja, um grupo inicial contendo todos os elementos é dividido em dois subgrupos, de tal forma que os elementos em um subgrupo estejam distantes dos elementos do outro subgrupo. Esses subgrupos são então divididos em subgrupos dissimilares e o processo continua até cada elemento formar um grupo. O método divisivo exige um maior número de iterações (que cresce exponencialmente), começará considerando todas as divisões dos elementos em dois grupos, com pelo menos um elemento em cada grupo.

Os métodos não-hierárquicos, ou por particionamento, foram desenvolvidos para agrupar elementos em  $K$  grupos, onde  $K$  é a quantidade de grupos definida previamente. Nem todos valores de  $K$  apresentam grupos satisfatórios, sendo assim, aplica-se o método várias vezes para diferentes valores de  $K$ , escolhendo os resultados que apresentem melhor interpretação dos grupos ou uma melhor representação gráfica (BUSSAB, MIAZAKI e ANDRADE, 1990).

A ideia central da maioria dos métodos por particionamento é escolher uma partição inicial dos elementos e, em seguida, alterar os membros dos grupos para obter-se a melhor partição (ANDERBERG, 1973).

Quando comparado com o método hierárquico, o método por particionamento é mais rápido porque não é necessário calcular e armazenar, durante o processamento, a matriz de similaridade.

Em geral, os métodos por particionamento diferem entre si pela maneira que constituem a melhor partição. Os métodos por particionamento mais conhecidos são o método *k-means* (k-médias) e o método *k-medoid* (k-medóides).

### 3 Metodologia

A trajetória metodológica desta pesquisa seguirá os seguintes passos:

- Análise estatística exploratória dos dados com o objetivo de descrever a informação presente nas Bases de Dados, caracterizar as escolas e conhecer os fatores educacionais presentes nos dados. Esta análise utilizou Medidas Estatísticas, Análise de Correlação Linear Simples e Análise de Conglomerados. Desta análise resultará um conjunto pré-selecionado de variáveis e grupos de escolas para análise de desempenho.
- A metodologia de avaliação do desempenho das escolas sob o ponto de vista da eficiência produtiva basear-se-á na Análise Envoltória de Dados aplicada a subconjuntos de escolas classificadas a partir da infraestrutura disponível.
- A aplicação de DEA permitirá a identificação das escolas de melhor desempenho e a determinação de ações e estratégias de melhoria para as escolas consideradas ineficientes.

Os passos apresentados permitiram alcançar os objetivos propostos.

## 4 Aplicação do Modelo DEA a Escolas Públicas do DF

### 4.1 Análise Exploratória dos Dados

#### 4.1.1 Informações Disponíveis

O Instituto Nacional de Estudos e Pesquisas Educacionais (INEP), é uma autarquia vinculada ao Ministério da Educação e tem como missão institucional produzir e disseminar informações educacionais. As estatísticas e avaliações produzidas pelo INEP visam fornecer os subsídios necessários para a formulação e implementação de políticas voltadas para a melhoria contínua da educação no País. Dentre as informações educacionais produzidas pelo INEP, destaca-se o Censo Escolar e os resultados da Prova Brasil.

Os dados da Secretaria de Educação - SE permitem determinar quais escolas pertencem a uma determinada Região Administrativa - RA e Coordenação Regional de Ensino-CRE. Além de saber se determinadas escolas são administradas pela SE.

#### 4.1.2 Dados

Os dados utilizados foram os microdados da Prova Brasil, o Censo Escolar, a Taxa de Rendimento das Escolas e dados da Secretaria da Educação do Distrito Federal.

As escolas selecionadas para o estudo de DEA foram as Escolas Públicas do Distrito Federal. Foram desconsideradas as escolas federais como o Colégio Militar e o Dom Pedro II porque possuem características diferentes das outras escolas públicas. Os alunos que ingressam nessas escolas serão obrigados a fazer um Vestibulinho que é uma espécie de vestibular. Nas escolas públicas do Distrito Federal o aluno tem direito à vaga e, para se matricular, basta o responsável ligar para 156 ou ir à escola que almeja. Se a escola não tem vaga, o aluno é encaminhado para outra escola.

Pelo site do INEP tem-se que, no Distrito Federal, 88,5% das escolas públicas são urbanas enquanto 11,5% são rurais como mostra a Tabela 1.

**Tabela 1: Número de escolas da rede pública do DF por situação urbana e rural segundo CRE.**

CRE	INSTITUIÇÕES EDUCACIONAIS		
	Urbana	Rural	Total
PLANO PILOTO - CRUZEIRO	103	0	103
GAMA	43	7	50
TAGUATINGA	62	4	66
BRAZLANDIA	18	12	30
SOBRADINHO	35	11	46
PLANALTINA	44	20	64
NUCLEO BANDEIRANTE	31	3	34
CEILANDIA	89	0	89
GUARA	25	0	25
SAMAMBAIA	40	0	40
SANTA MARIA	26	0	26
PARANOÁ	17	14	31
SAO SEBASTIAO	19	4	23
RECANTO DAS EMAS	24	0	24
<b>Total</b>	<b>576</b>	<b>75</b>	<b>651</b>

FONTE: CENSO ESCOLAR - SE/DF

Tem-se 506 escolas de ensino fundamental, 338 escolas oferecem Ensino Fundamental I (1º a 5º ano) e 168 escolas o Ensino Fundamental II (6º a 9º ano) como mostra a Tabela 2.

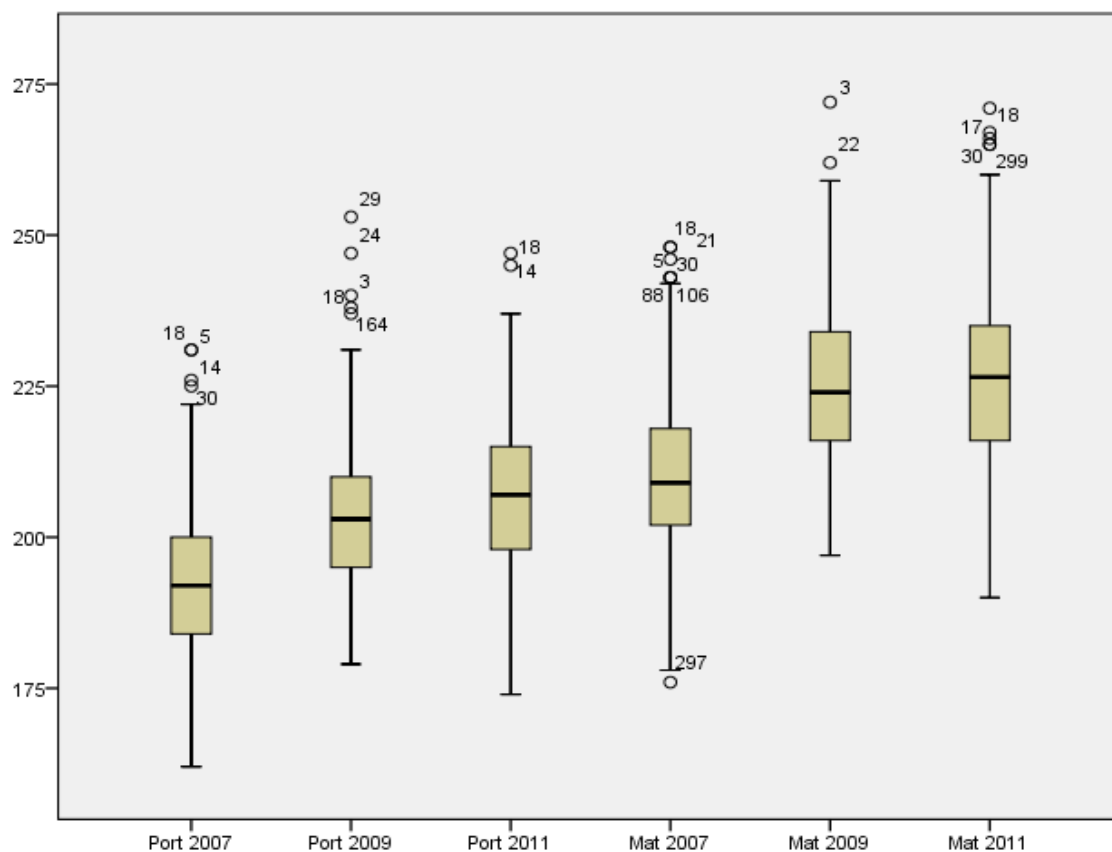
**Tabela 2: Número de escola no ensino fundamental por nível de ensino segundo a CRE em 2013.**

CRE	1º a 5º ano	6º a 9º ano
Núcleo Bandeirante	20	11
Brazlândia	15	8
Ceilândia	56	24
Gama	29	16
Guara	11	8
Paranoá	13	7
Planaltina	34	18
Plano Piloto	41	16
Recanto das Emas	16	11
Samambaia	22	12
Santa Maria	13	8
São Sebastião	12	5
Sobradinho	23	11
Taguatinga	33	13
<b>Total Geral</b>	<b>338</b>	<b>168</b>

Fonte: INEP – CENSO ESCOLAR 2013

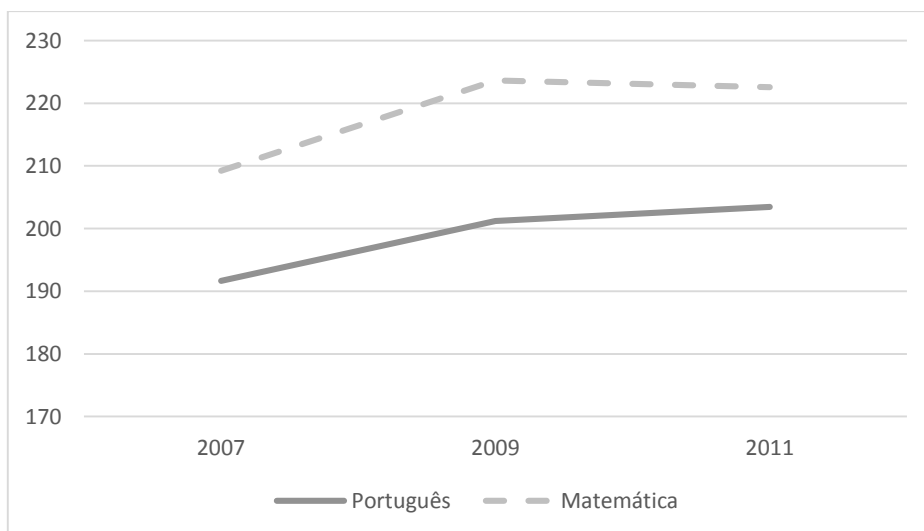
Da base de dados da Prova Brasil, foram extraídas as notas em Português e Matemática dos anos de 2007, 2009 e 2011. Aparentemente, as notas de matemática são maiores do que as de português, porém, essas notas estão em escalas diferentes. No caso do português começa com 125 e vai a 350, enquanto a matemática começa com 125 e vai a 425 como mostram nas Figuras 9 e 10.

**Figura 9: Nota da Prova Brasil do 5º ano da Escola Pública do Distrito Federal.**



Fonte: INEP – Prova Brasil

**Figura 10: Média ponderada da Prova Brasil do 5º ano escola pública do GDF**



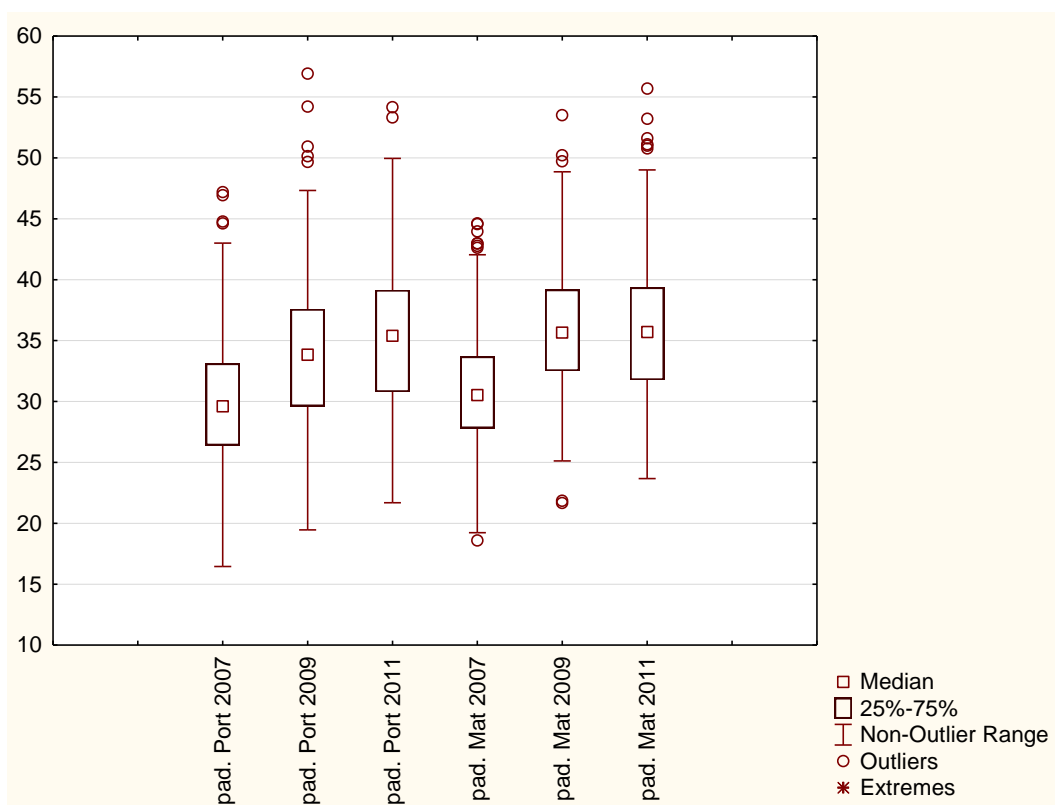
Fonte: INEP – Prova Brasil

Com a média ponderada da Prova Brasil do 5º ano na Figura 10 tiveram um aumento em 2007 para 2009 nas duas matérias. Porém, na matemática houve um decrescimento baixo de 2009 a 2011, enquanto em português houve um leve crescimento neste período. De acordo com a escala de desempenho do INEP, os alunos estão no nível 3 em português e nível 5 em matemática (informação mais detalhado no anexo). Além disso, não podemos dizer que os alunos sabem mais matemática do que português porque as escalas são diferentes. No caso da matemática a prova começa com 125 e o seu máximo é 425 enquanto português começa com 125 e o máximo é 325. Em razão disso, o gráfico não é possível dizer se um é melhor do que o outro. Para viabilizar a comparação entre as notas fez-se uma padronização da seguinte forma:

$$Valorpadronizado = \frac{100 \cdot (média\ ponderada - nota\ mínima)}{(nota\ máxima - nota\ mínima)}$$

Com isso, os valores passam a ser de 0 a 100 e poderá comparar português e matemática, conforme pode se observar na Figura 11.

**Figura 11: Padronização da nota do 5º ano da Prova Brasil da escola pública do DF.**



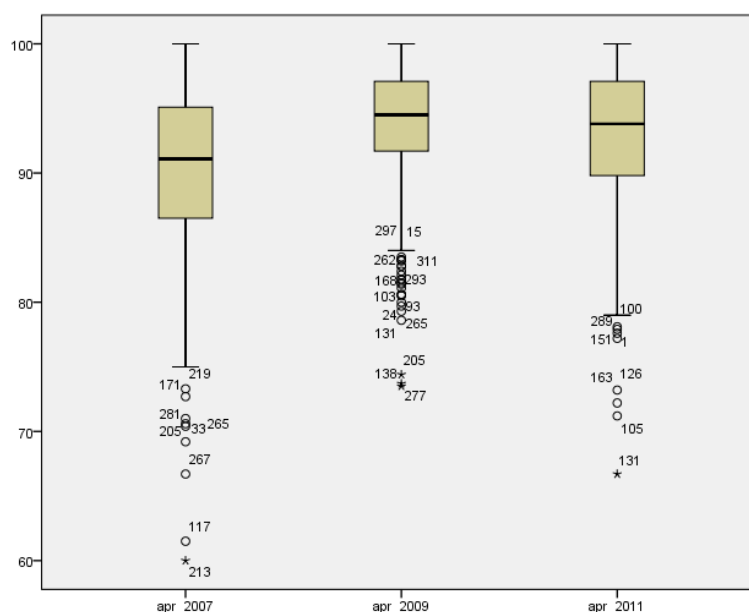
Fonte: INEP – Prova Brasil

Com a figura 11 a nota da Prova Brasil do 5º ano tem-se desempenhos próximos em português e matemática. Por outro lado, os alunos possuem médias inferior a 40 e se utilizasse o critério das escolas públicas para a aprovação que é de 50 da nota, implicaria maioria dos alunos seriam reprovados.

A taxa de aprovação é outro indicador de qualidade e neste caso foram considerados apenas os anos que tiveram a Prova Brasil, a Figura 12 mostra que a maioria dos alunos são aprovados porque a mediana da taxa é superior a 90% nos anos considerados. Ou seja, a maioria dos alunos foram aprovados sem ter uma base sólida em matemática e português.

Tendo um dos motivos o regimento escolar pode aprovar o aluno mesmo que o professor seja discordante com a sua aprovação.

**Figura 12: Taxa de aprovação 5º ano escola pública do Distrito Federal.**



Fonte: INEP – CENSO ESCOLAR 2013

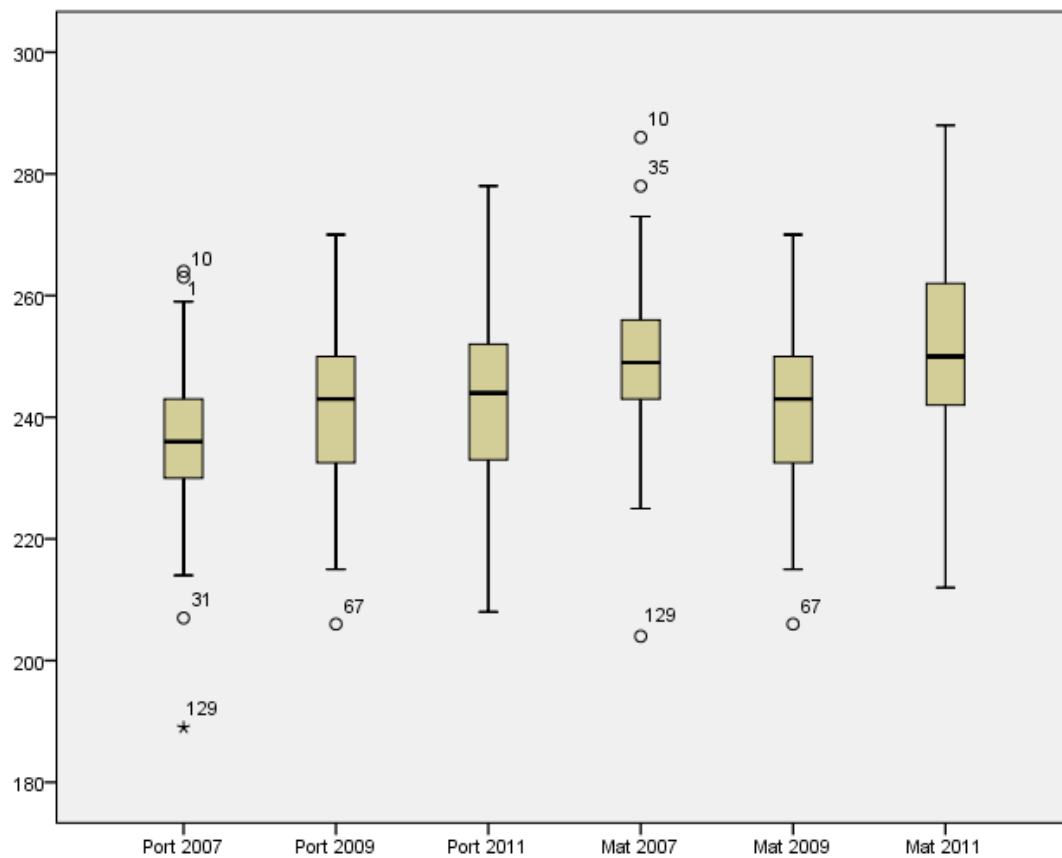
Pelas Figuras 11 e 12 de taxa de aprovação do 5º ano a maioria das escolas tiveram a taxa de aprovação acima de 80% que contradiz com a realidade da Prova Brasil padronizada que mostrou notas abaixo de 50 pontos.

De forma análoga, padronizou-se as notas do 9º ano.

Com a média ponderada da nota da Prova Brasil do 9º ano tem-se um decréscimo de 2007 para 2009 em matemática e um crescimento de 2009 para 2011, enquanto em português houve um crescimento de 2007 a 2009 e um leve crescimento de 2009 para 2010. Pela escala de desempenho do INEP, os alunos estão no nível 5 em português e nível 5 e 6 em matemática (informação mais detalhada na Tabela 13 do Anexo) como mostram as Figuras 13 e 14.

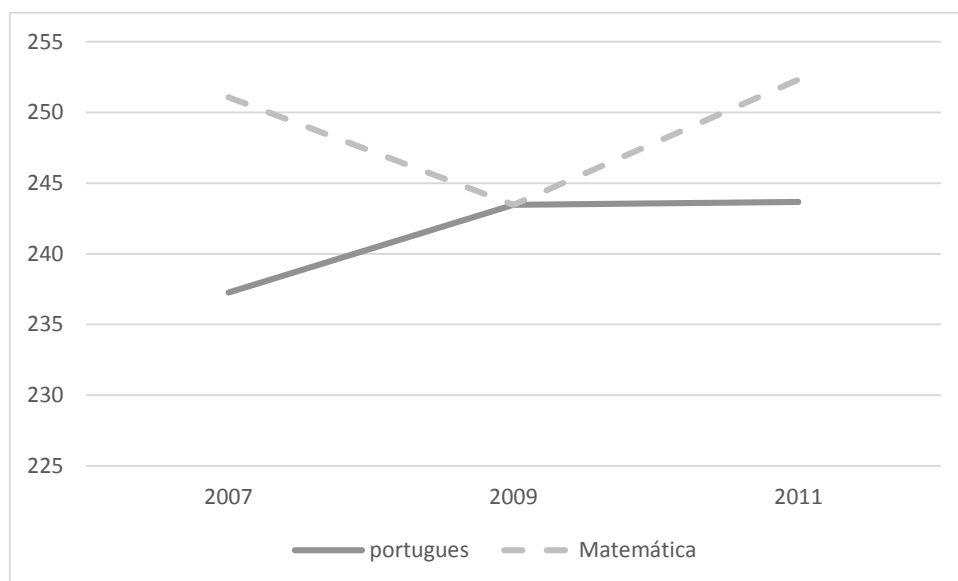


**Figura 13: Nota da Prova Brasil do 9º ano da Escola Pública do Distrito Federal**



Fonte: INEP – Prova Brasil

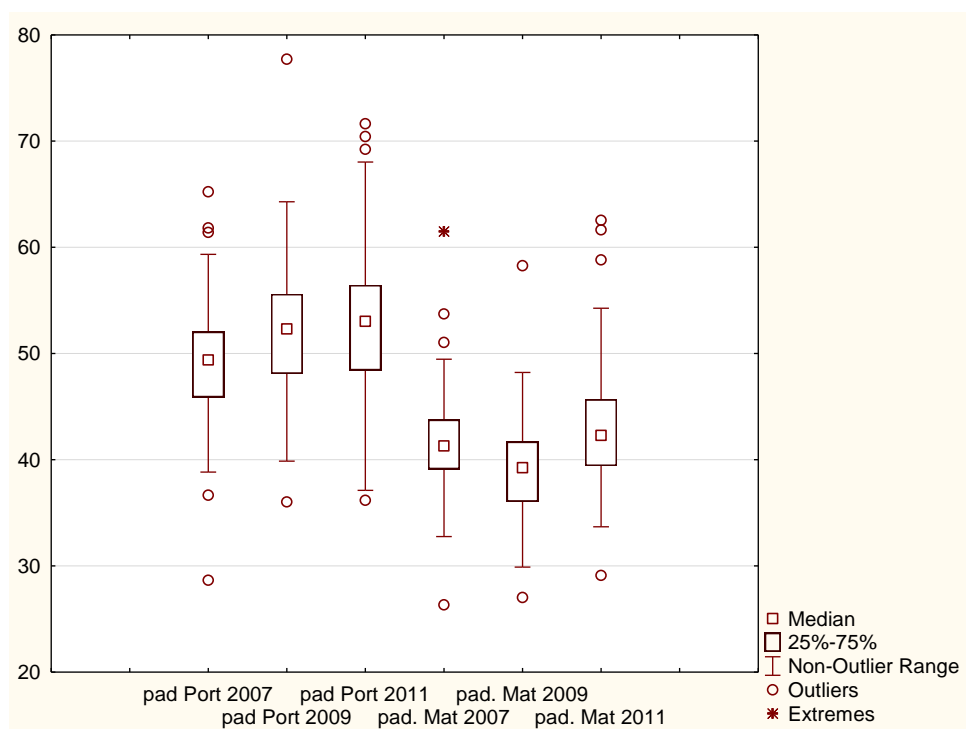
**Figura 14: Média ponderada da prova brasil, 9º ano escola pública do GDF**



Fonte: INEP – Prova Brasil

Como não há possibilidade de comparar, na forma análoga ao 5º ano, fez-se a padronização das notas da Prova Brasil, como mostra a Figura 15 abaixo:

**Figura 15: Padronização da nota do 9º ano da Prova Brasil das escolas públicas do DF.**

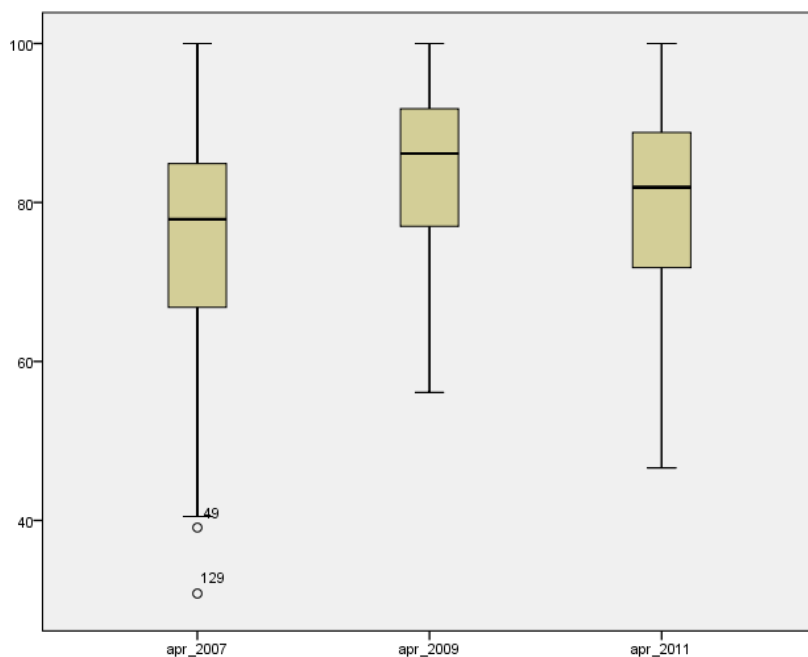


Fonte: INEP – Prova Brasil

Este gráfico mostra que os alunos tiveram melhor desempenho em português do que em matemática. Além disso, a nota de português foi acima de 50 enquanto a matemática foi pouco abaixo de 50 pontos.

Por outro lado, o 9º ano mostra que há mais alunos retidos na série, porém ainda mostra que os alunos passam de série sem dominar o conteúdo proposto de acordo com as Figuras 15 e 16.

**Figura 16: Taxa de aprovação 9º ano escola pública do Distrito Federal.**



Fonte: INEP

Também foram analisados os recursos das escolas, número de docentes e titulação, entre outros, para verificar sua influência da qualidade do ensino. E um fato curioso é que entre os professores do 1º a 5º ano, a maioria possui pós-graduação, enquanto no 6º a 9º ano a maioria não possui pós-graduação, como mostra a Tabela 3.

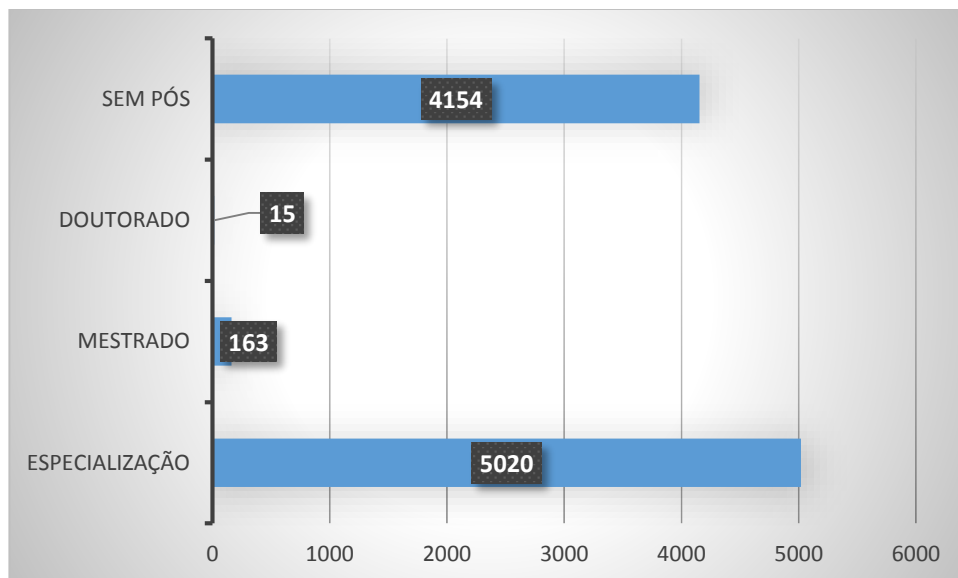
**Tabela 3: Titulação dos professores das escolas públicas do GDF em 2013.**

	Especialização	Mestrado	Doutorado	Sem pós	Total
1º a 5º ano	4997	162	15	4145	9319
6º a 9ºano	3130	243	30	4781	8184
Total	8127	405	45	8926	17503

Fonte: INEP – CENSO ESCOLAR 2013

Ou seja, no 1º a 5º anos os 53,6% dos professores possuem uma especialização, 1,7% mestrado, 0,2 % doutorado e 44,5% não possuem pós-graduação.

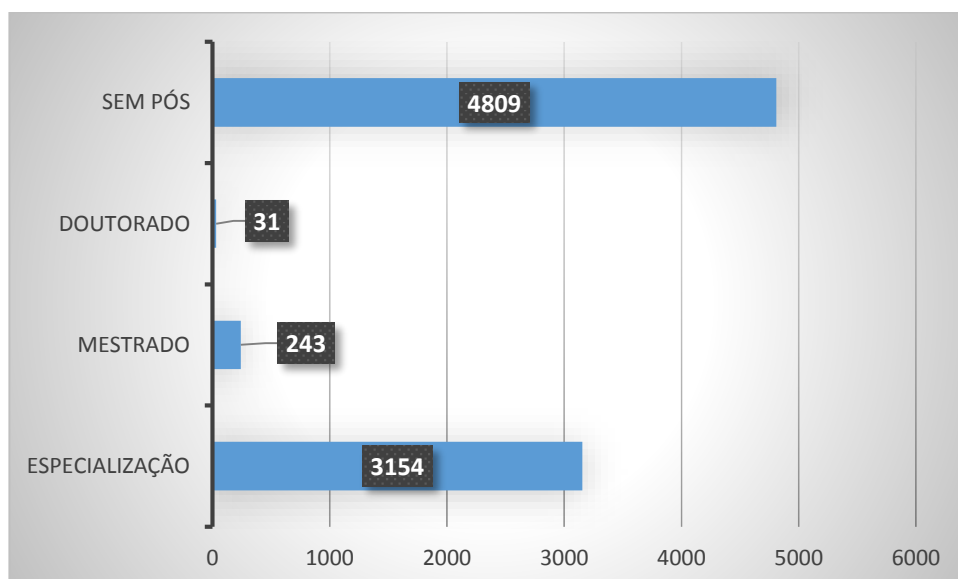
**Figura 17: Titulação dos docentes do 1º a 5º ano em 2013.**



Fonte: CENSO ESCOLAR 2013.

Por outro lado, entre o 6º a 9º anos há mais professores sem pós-graduação do que com especialização, isto que, 38,2% possuem especialização, 3,0% mestrado, 0,4% doutorado e 58,4% não possuem pós-graduação.

**Figura 18: Título do docente 6º a 9º ano em 2013.**



Fonte: CENSO ESCOLAR 2013.

Outras informações sobre os recursos da escola referem-se a instalações administrativas, didático-pedagógicas e equipamentos disponíveis a professores e alunos, conforme a Tabela 4. A informação sobre a disponibilidade destes recursos está no Censo Escolar e permitiu a construção de indicadores que qualificam a infraestrutura da escola.

**Tabela 4: Recursos e equipamentos pedagógicos**

Código	Significado
ID_LABORATORIO_INFORMATICA	LABORATORIO INFORMATICA
ID_LABORATORIO_CIENCIAS	LABORATORIO CIENCIAS
ID_SALA_ATENDIMENTO_ESPECIAL	SALA ATENDIMENTO ESPECIAL
Quadra	QUADRA
ID_BIBLIOTECA	BIBLIOTECA
ID_PARQUE_INFANTIL	PARQUE INFANTIL
ID_DEPENDENCIAS_PNE	DEPENDENCIAS PNE
ID_REFEITORIO	REFEITORIO
ID_DESPENSA	DESPENSA
ID_ALMOXARIFADO	ALMOXARIFADO
Pátio	PATIO
ID_AREA_VERDE	AREA VERDE
ID_AEE	Atendimento Educacional Especializado
NUM_SALAS_EXISTENTES	Número de Salas Existentes
NUM_SALAS_UTILIZADAS	Número de Salas Utilizadas
NUM_EQUIP_TV	Número EQUIPAMENTO DE TV
NUM_EQUIP_DVD	Número EQUIPAMENTO DVD
NUM_EQUIP_PARABOLICA	Número de EQUIPAMENTO PARABOLICA
NUM_EQUIP_COPIADORA	Número EQUIPAMENTO COPIADORA
NUM_EQUIP_IMPRESSORA	Número EQUIPAMENTO IMPRESSORA
NUM_EQUIP_SOM	Número EQUIPAMENTO SOM
NUM_EQUIP_MULTIMIDIA	Número EQUIPAMENTO MULTIMIDIA
NUM_EQUIP_FOTO	Número EQUIPAMENTO FOTO
NUM_COMPUTADORES	Número DE COMPUTADORES
ID_INTERNET	INTERNET

Fonte: INEP – Censo Escolar 2013

#### 4.1.3 Seleção das escolas

Foram construídos indicadores para resumir os aspectos de um grupo de variáveis relacionadas a alguma característica da escola. A maioria das variáveis que apresentam características das escolas é do tipo categórica e possui uma ordem natural: Se a escola possui uma determinada dependência, se a escola possui um grande número de determinados equipamentos e se os equipamentos são devidamente utilizados na escola. É intuitivo presumir que alunos de uma escola com uma boa estrutura com relação a dependências, equipamentos e utilização dos recursos disponíveis na instituição têm melhores condições de aprendizagem e, em consequência, tendem a obter melhores desempenhos na Prova Brasil.

Os indicadores construídos têm como objetivo quantificar a qualidade da escola quanto a existência de dependências, quantidade e utilização de equipamentos e recursos disponíveis nas instituições, para testar a viabilidade de avaliá-las em um mesmo modelo de eficiência.

Inicialmente há dois tipos de variáveis que compõem os indicadores: Qualitativas dicotômicas do tipo “Sim” ou “Não” (variáveis de existência de dependências e utilização de recursos) e variáveis discretas (quantidade de equipamentos).

##### **Indicador da existência de instalações**

Escolhido um conjunto de instalações, adotou-se como indicador da existência de instalações a proporção dessas instalações que existem na escola.

As instalações consideradas na construção do indicador foram:

- Laboratório de informática;
- Laboratório de ciências;
- Sala de atendimento especial;
- Quadra de esportes;
- Biblioteca;
- Dependência de PNE;
- Refeitório;
- Despensa;
- Almoxarifado;
- Pátio;
- Área verde;
- Atendimento Educacional Especializado (AEE)

Instalações relativas a outras dependências não foram consideradas porque quase todas as escolas possuem (como sala da direção, sala de professores, secretaria, banheiro, entre outros), ou porque nenhuma escola possui a exemplo de banheiro externo a escola.

Além disso, foi construído a variável televisão por sala de aula (*TV/SA*) porque a televisão atual substitui vários aparelhos como som, kit multimídia, DVD. Sendo um grande recurso que o professor pode acoplar no computador para fazer diversos, trabalhos como gráfico de funções na matemática ou apresentar slides. Sem esquecer, alunos podem fazer slide em pdf e acoplar direto na televisão evitando diversos vírus.

Outra variável considerada foi Número de Computadores vezes Números de Salas sobre Números de alunos (*Comp*), por ser um recurso importante aos alunos, pois nem sempre têm

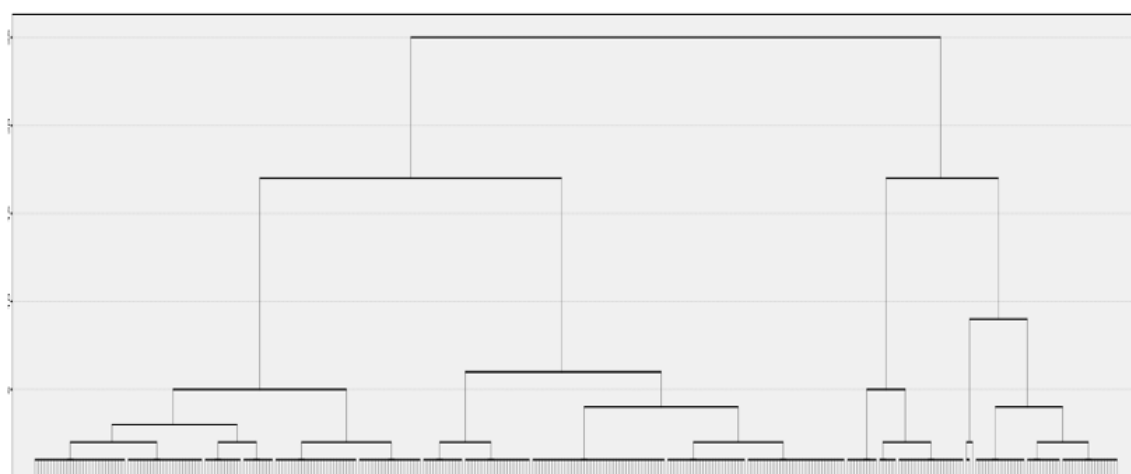
computadores em suas casas. Por outro lado, o professor poderá ensinar aos alunos utilizar aplicativos que facilitam o aprendizado por exemplo o Winplot, Excel.

A partir da formação dos indicadores, que representam características estruturais das instituições, as escolas foram separadas em grupos cujos elementos apresentam indicadores com características semelhantes. Neste sentido, escolas que apresentam indicadores da infraestrutura muito distintos não devem ser comparadas. Para classificar as escolas em grupos homogêneos, a Análise de Conglomerados foi utilizada. Algumas escolas apresentam indicadores com valores elevados e outras com valores baixos. Por este motivo, o método de análise de conglomerados utilizado foi o de variância mínima (*Ward*), uma técnica hierárquica aglomerativa que consiste em tratar cada elemento inicialmente como um conglomerado. As combinações de escolas que produzem as menores variâncias são agrupadas em um conglomerado e, em seguida, são calculadas as novas variâncias dos conglomerados formados. A técnica se repete até que um conglomerado contendo todas as escolas seja formado. O objetivo deste método é obter um número de grupos que produzam as menores variâncias internas. Os indicadores possuem escalas diferenciados, portanto para a aplicação da técnica de formação de grupos os indicadores serão transformados de forma que todos possuam média 0 (zero) e variância 1 (um). Desta forma as escalas dos indicadores não exercem influência na aplicação da técnica. A medida de similaridade entre os elementos utilizada neste método é a distância euclidiana. Com isso, será construído um dendograma com as variáveis Renda da Região Administrativa (*Renda*), *Instalações*, *TV/SA* e *Comp*.

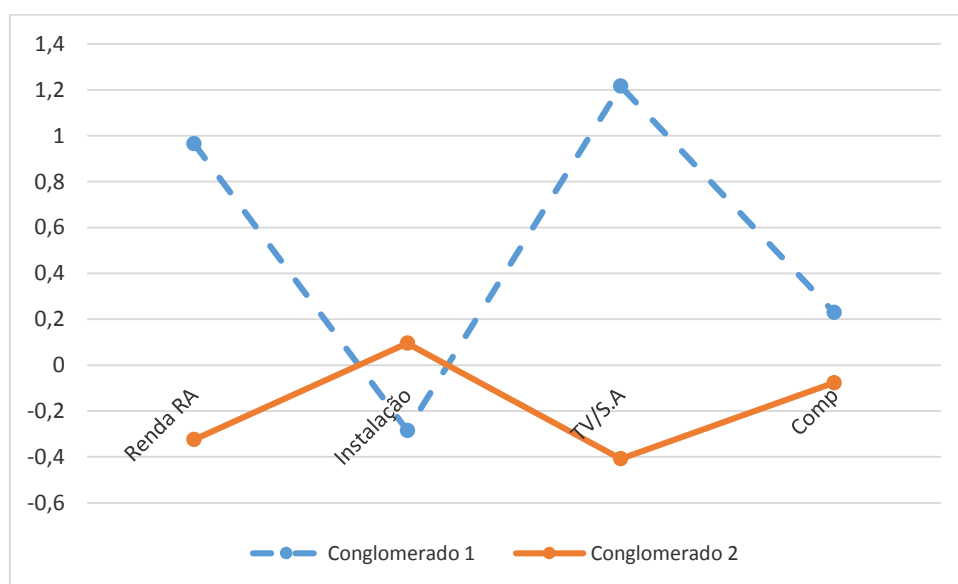
#### 4.1.3.1 Seleção das escolas do 5º ano.

Com as variáveis propostas aplicou-se Análise de Conglomerados nas 388 escolas do 5º ano. Observou-se duas estruturas consistentes de grupos; uma delas com quatro grupos, basicamente determinados, cada um deles, por uma das variáveis utilizadas; e outra em dois grupos separando escolas bem equipadas em regiões de renda mais alta e escolas menos equipadas em regiões de renda mais baixa. Considerando os objetivos do trabalho, adotou-se a classificação em dois grupos. Assim, o primeiro (*conglomerado 1*) tem 85 escolas e se caracteriza por escolas melhores equipadas em regiões de renda mais alta, enquanto o segundo (*conglomerado 2*), com 253 escolas em regiões de renda mais baixa e piores condições de infraestrutura, como mostram as figuras 19 e 20.

**Figura 19: Dendograma 5º ano**



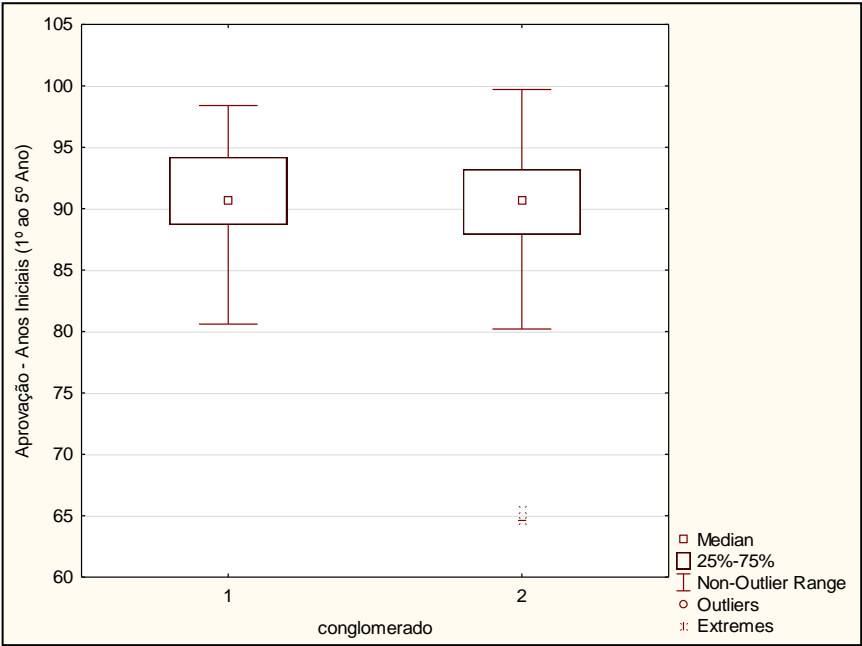
**Figura 20: Média padronizada das variáveis separados por conglomerado 1 e 2.**



As escolas do 1º a 5º ano que estão localizados em regiões mais ricas e com melhores infraestrutura possuem taxa de aprovação praticamente idêntica às escolas que estão localizados em regiões mais pobres e com pior infraestrutura, conforme se vê na Figura 21.

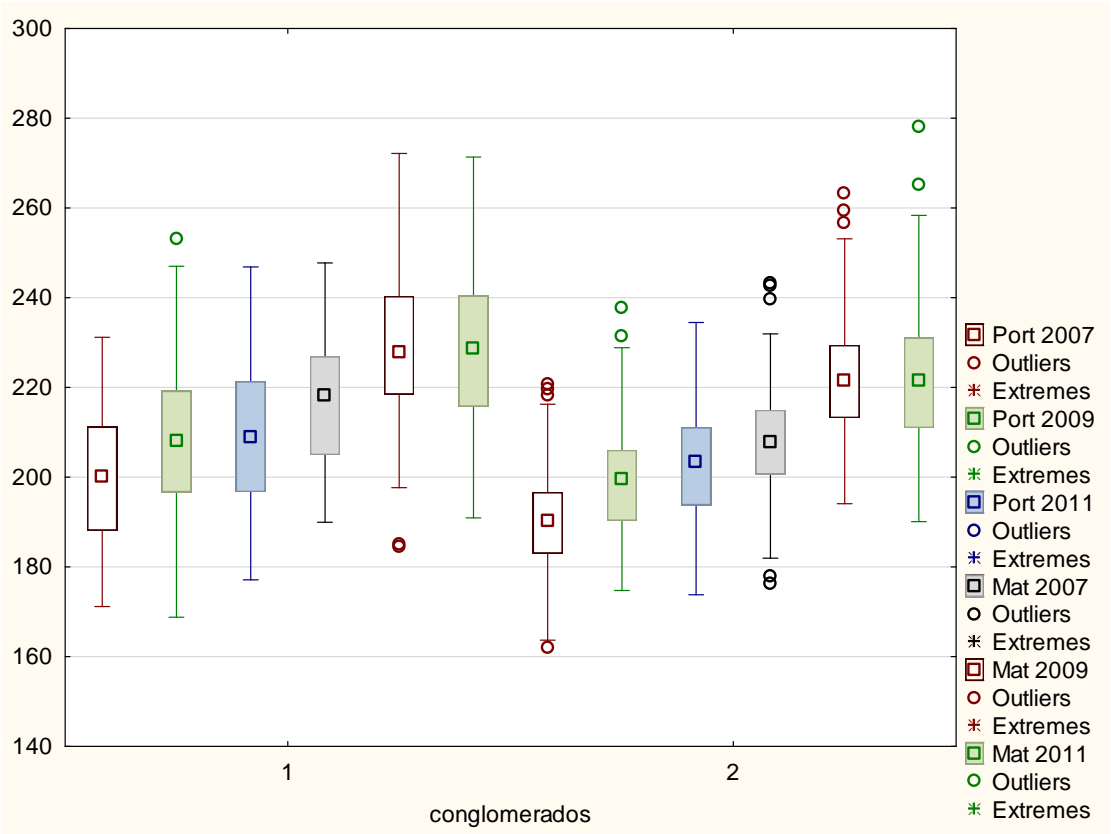


Figura 21: Box Plot de taxa de aprovação segundo conglomerado 1 e 2.



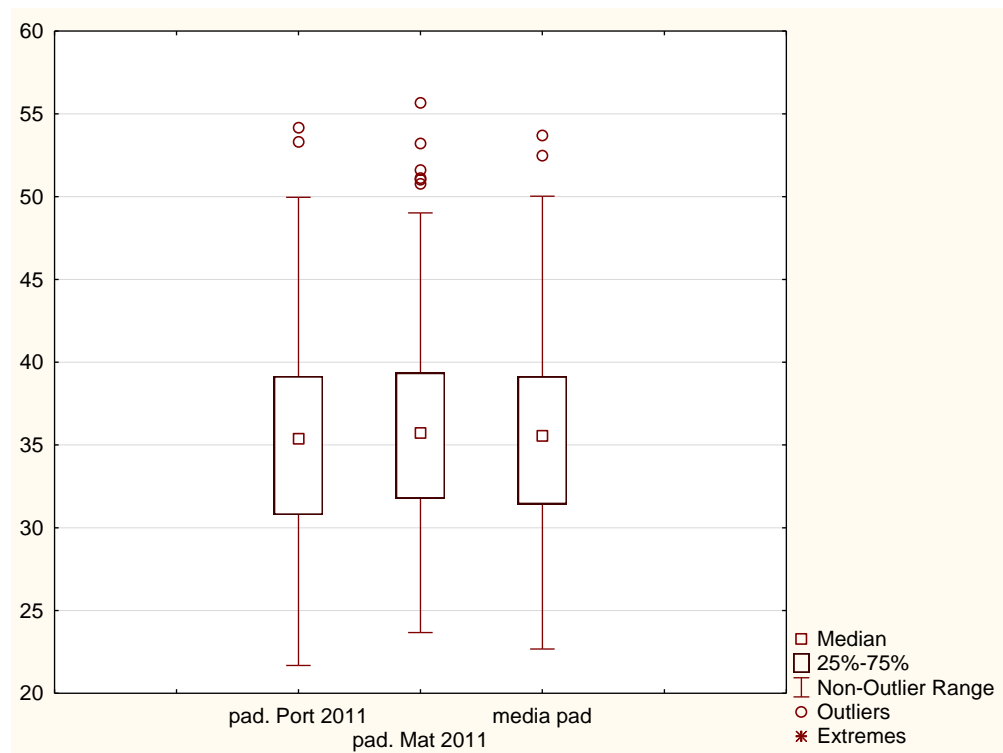
Por outro lado, a figura 22 mostra que no conglomerado 1 os alunos possuem melhores rendimentos na Prova Brasil que no conglomerado 2.

Figura 22: Nota da Prova Brasil segundo conglomerados 1 e 2.



Como foi dito acima, trabalha-se com a padronização da nota da Prova Brasil e pelo gráfico de Box-Plot tem-se a nota de português 2011, matemática 2011 e média dessas duas notas. Como mostra a Figura 23 as três notas têm o mesmo comportamento e com isso será utilizado na análise de eficiência a média da Prova Brasil padronizada como descritora de desempenho na Prova Brasil.

**Figura 23: Box Plot da Prova Brasil de Português, Matemática e média dos dois em 2011.**

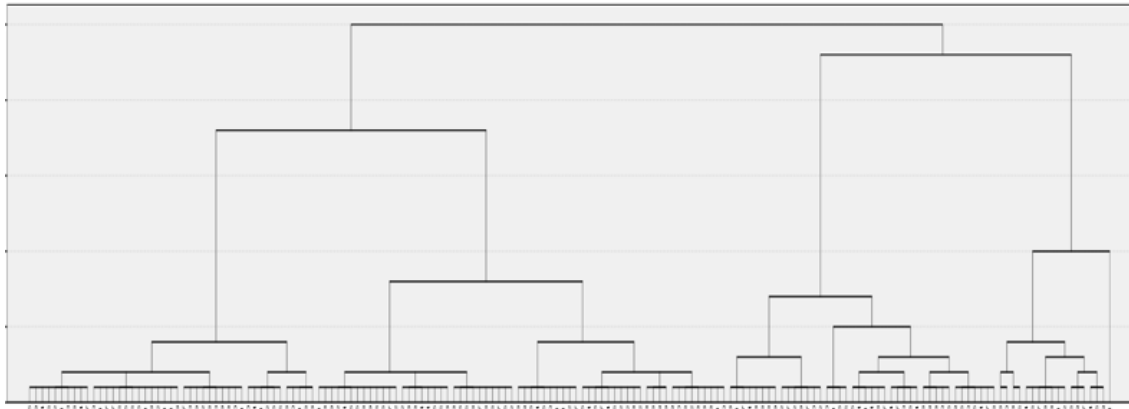


A análise identificou dois conglomerados no 5º ano com características diferentes de renda e infraestrutura. Assim a aplicação de DEA se fez dentro de cada Conglomerado.

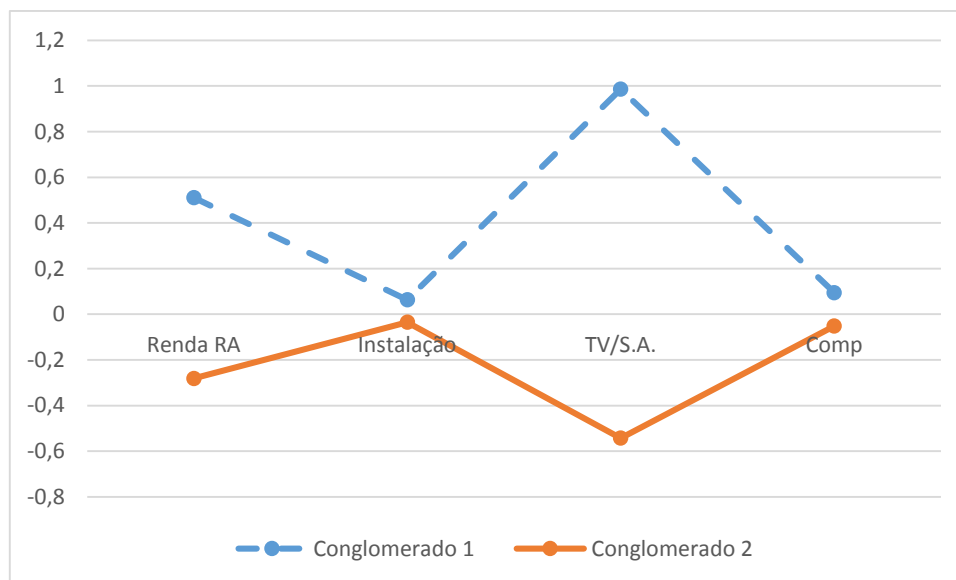
#### 4.1.3.2 Seleção das escolas 9º ano.

Com as mesmas variáveis que são *Renda da Região Administrativa*, *instalações*, *TV/SA* e *Comp*, novamente identificou-se dois conglomerados com as mesmas características encontradas no 5º ano conforme as Figuras 24 e 25. A análise de conglomerados sugere a formação de dois conglomerados, o primeiro (conglomerado 1) tem 60 escolas e se caracteriza por escolas melhores equipadas em regiões de renda mais alta, enquanto o segundo (conglomerado 2), com 109 escolas em regiões de renda mais baixa e piores condições de infraestrutura.

**Figura 24: Dendograma 9º ano**



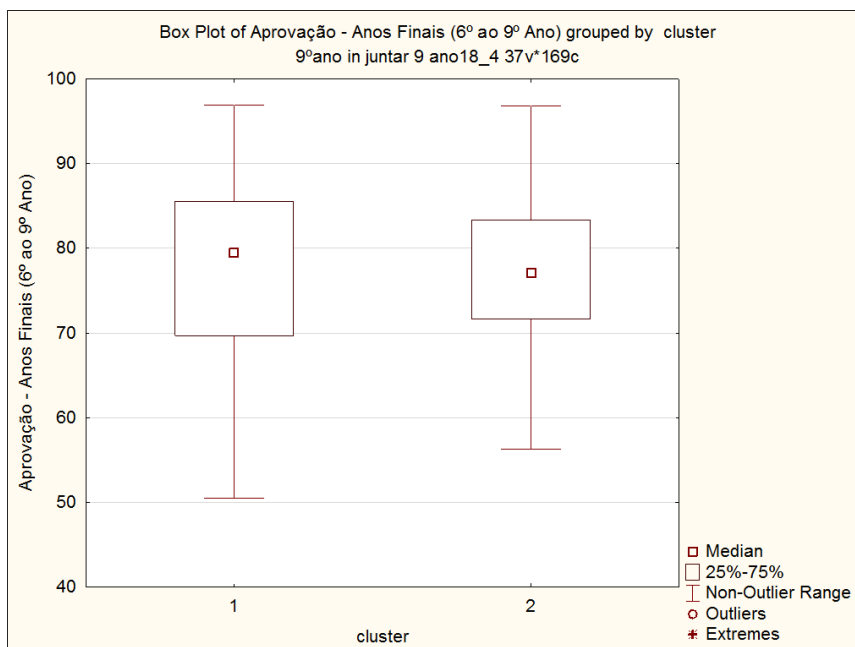
**Figura 25: Média padronizada das variáveis separados por conglomerado 1 e 2.**



A Análise de Conglomerados identificou dois conglomerados no 9º ano com características diferentes de renda e infraestrutura. Assim a aplicação de DEA se fez dentro de cada conglomerado.

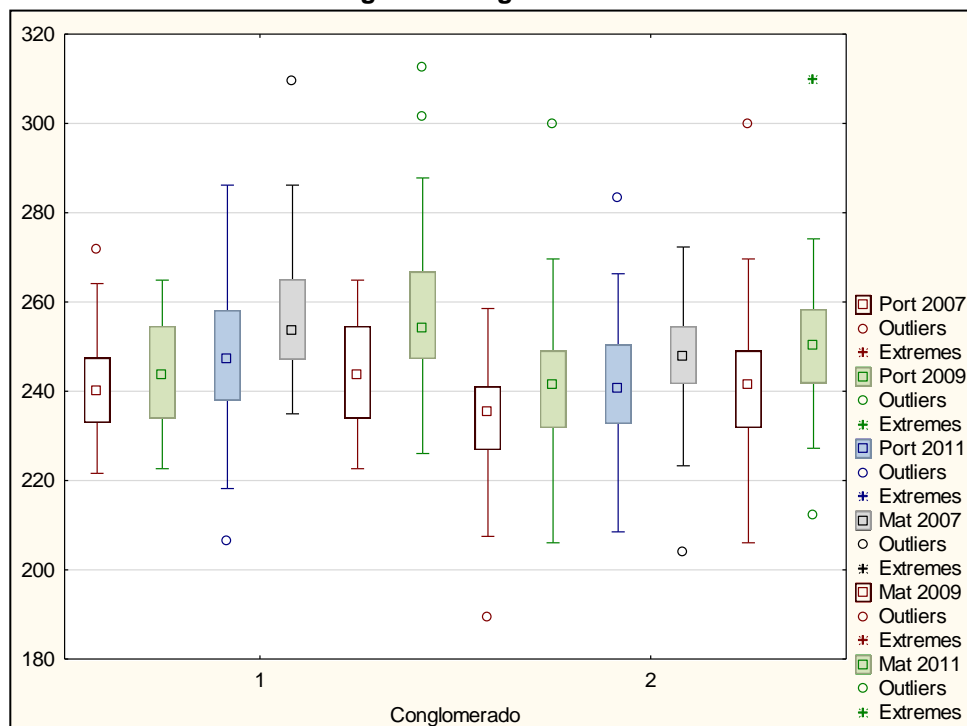
Pela Figura 26, tem-se a divisão de dois grandes grupos que chamaremos de conglomerado 1 e conglomerado 2. Na divisão os dois conglomerados são praticamente idênticos em termos da taxa de aprovação dos alunos do 6º ao 9º ano.

**Figura 26: Box Plot de taxa de aprovação segundo conglomerado 1 e 2.**



Por outro lado, a Figura 27 mostra que no conglomerado 1 os alunos possuem melhores rendimentos na Prova Brasil que no conglomerado 2.

**Figura 27: Nota da Prova Brasil segundo conglomerado 1 e 2.**



## 4.2 Modelo DEA

Para construir o modelo DEA, aplicou-se o procedimento de seleção de variáveis de Norman & Stoker a cada um dos quatro conglomerados identificados. Além disso, para medir a eficiência que é uma razão de produto sobre insumo tem-se que identificar quais elementos são insumos, que são variáveis utilizadas nas escolas para chegar a um determinado resultado. E quais elementos são o produto que são as variáveis resultados. Neste caso os insumos são *instalações, comp, TV/SA*, número de professores, titulação dos professores e os produtos são resultado da Prova Brasil, Taxa de Aprovação, número de alunos.

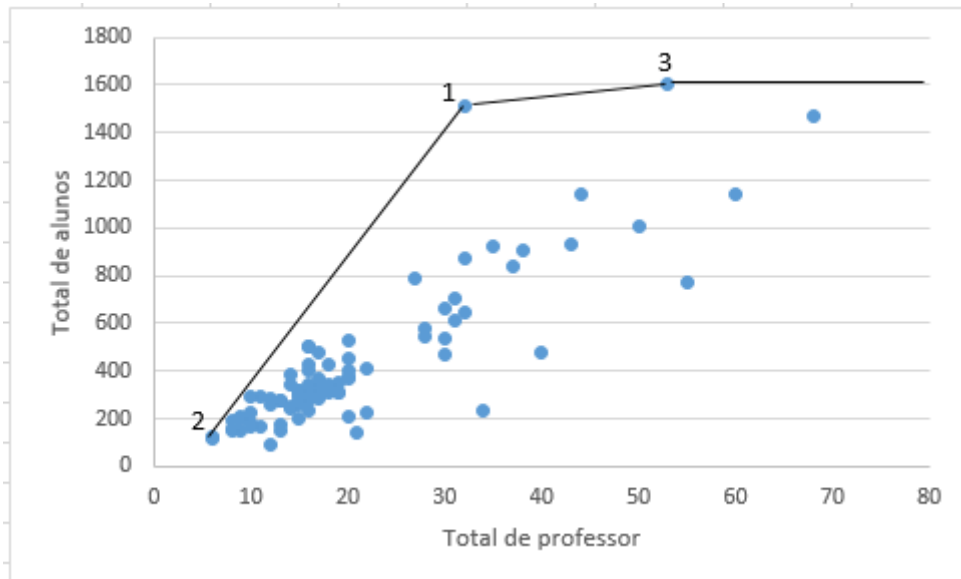
### 4.2.1 Modelo DEA 5º ano conglomerado 1.

Seguindo a metodologia de Norman & Stoker o primeiro modelo consiste em maior correlação é total de alunos e total de professor. Com isso, será construído o primeiro modelo do DEA1, porque a correlação de total de alunos e total de professor é 0,86 conforme a Tabela 14 do Anexo.

$$DEA1 = \frac{\text{Total de alunos}}{\text{Total de professor}}$$

A Figura 28 apresenta a fronteira de eficiência, enquanto a Tabela 5 identifica as escolas eficientes obtidas com modelo DEA1.

**Figura 28: DEA1, 5º ano, conglomerado 1**



**Tabela 5: Escolas eficientes do DEA1.**

Eficientes	Total de professores	Total de alunos
1 - CEF 02 DE Brazlândia	53	1600
2 - EC AGUILHADA	6	128
3 - CEF 02 DA ESTRUTURAL	32	1509

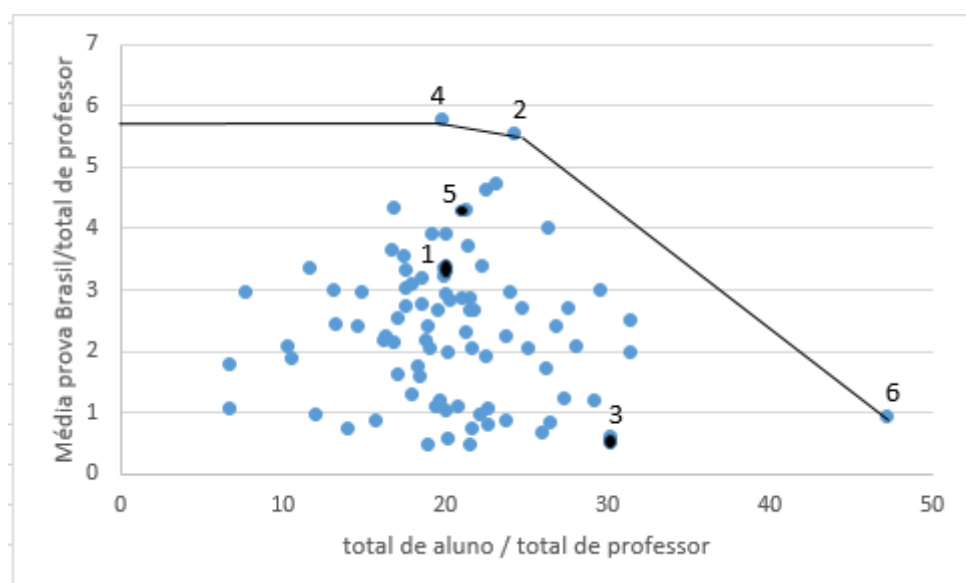
A EC AGUILHADA é eficiente em função do seu porte. Por ter o menor número de professores e menor número de aluno do sistema ela não pode ser escrita como combinação convexa das demais escolas, sendo assim, considerada eficiente em razão da adoção do modelo DEA de retorno variáveis a escala. O CEF 02 DA ESTRUTURAL domina a relação total de alunos por total de professor. O CEF 02 DE Brazlândia tem esta produtividade menor que o CEF 02 DA ESTRUTURAL, porém é maior em relação aos números de alunos, não podendo assim ser descrito pelas demais escolas eficientes identificadas.

Analisando a associação entre a medida de eficiência construída pelo DEA1 e as demais variáveis observou-se uma correlação negativa entre DEA 1 e a média da Prova Brasil ( $r = -0,29$ , conforme Tabela 14 do anexo). Com isso, seguindo o procedimento de Norman & Stoker, a variável média da Prova Brasil foi incluída a função de desempenho resultando o modelo DEA2:

$$DEA2 = \frac{\text{Total de alunos} + \text{média da Prova Brasil}}{\text{Total de professor}}$$

Com esse modelo três novas escolas foram consideradas eficientes perfazendo um total de seis escolas eficientes no modelo DEA2, conforme a Figura 29 e a Tabela 6.

**Figura 29: DEA 2, Conglomerado 1, 5º ano.**



**Tabela 6: Escolas eficientes DEA 2, 5º ano, Conglomerado 1.**

Eficientes	Prof. total	Total alunos	Prova Brasil.
1 - EC 305 SUL	16	319	53,69
2 - EC 411 NORTE	8	194	44,51
3 - CEF 02 DE Brazlândia	53	1600	32,86
4 - EC KANEGAE	6	119	34,65
5 - EC AGUILHADA	6	128	25,76
6 - CEF 02 DA ESTRUTURAL	32	1509	30,61

A Figura 28 e a Figura 29 têm estruturas totalmente diferentes. A primeira Figura apresenta uma relação produto por insumo tradicional em modelos de produção tendo no eixo da abscissa o número de professores (insumo) e na ordenada o número de alunos matriculados (produto). O segundo gráfico é uma estrutura de produtividades parciais, tendo no eixo da abscissa a razão entre total de alunos e total de professor, ou seja, uma razão de produto por insumo. No eixo da ordenada tem-se a razão entre a nota média na Prova Brasil e o total de professor que também é outra razão de produto por insumo. As escolas consideradas eficientes no modelo DEA1 continuam eficientes no modelo DEA2 por características da DEA. Além dessas, o modelo DEA2 identificou três outras escolas eficientes que são EC 305 SUL (que domina a nota da Prova Brasil), a EC KANEGAE (que domina a relação Prova Brasil por Professor) e a EC 411 NORTE (que não domina nenhuma das produtividades parciais mas tem melhores indicadores que qualquer combinação convexa das demais escolas eficientes).

A variável taxa de aprovação possui a maior correlação com DEA2 em relação aos dados que não foram utilizados. Construiu-se o DEA3 que é:

$$DEA3 = \frac{\text{Total de alunos} + \text{média da Prova Brasil} + \text{Aprovação}}{\text{Total de Professores}}$$

Com esse modelo 78 escolas das 85 atribuíram pesos insignificante na variável Prova Brasil e pesos máximos para a variável Taxa de Aprovação, inviabilizando a proposta de usar as notas da Prova Brasil como medida da qualidade do ensino. Isto é consequência da taxa de aprovação não diferenciar as escolas, uma vez que todas elas têm nesta variável um valor muito alto. Com isso, adotou-se o modelo DEA2 para a análise de eficiência das escolas do 5º ano, conglomerado 1.

$$DEA2 = \frac{\text{Total de alunos} + \text{média da Prova Brasil}}{\text{Total de professores}}$$

#### 4.2.2 Modelo DEA – 5º ano Conglomerado 2 e 9º ano Conglomerados 1 e 2.

O mesmo critério de seleção de variáveis foi aplicado ao 5º ano Conglomerado 2, 9º ano Conglomerado 1 e 9º ano Conglomerado 2 e obteve-se o mesmo modelo. Assim o modelo DEA 2 foi o modelo adotado para análise de eficiência em ambos os grupos (conglomerado 1 e conglomerado 2) nos dois níveis de ensino (Ensino Fundamental I e Ensino Fundamental II).



## 5 Aplicação dos Modelos DEA

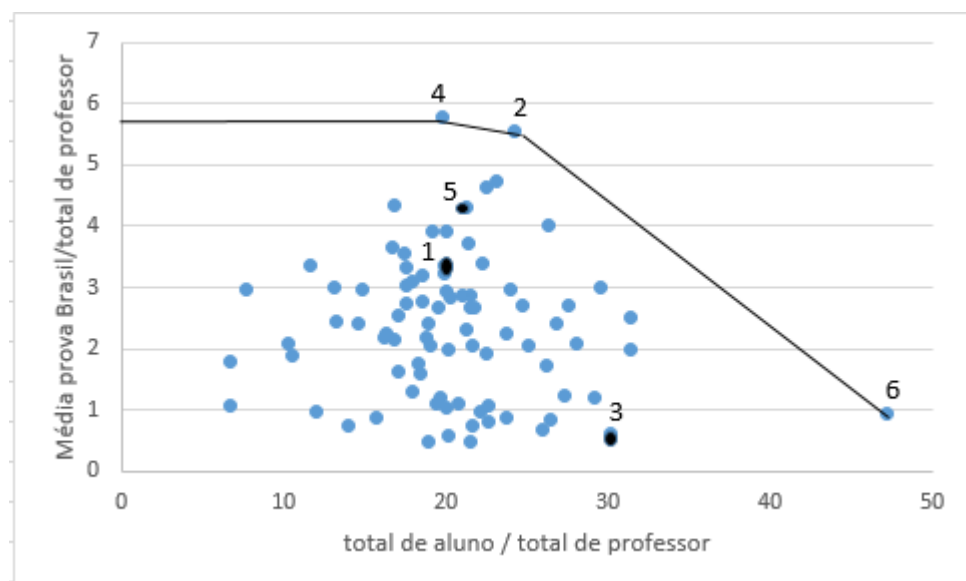
### 5.1 5º ano conglomerado 1

Aplicando-se o modelo DEA2 às escolas do 5º ano conglomerado1 identificou-se seis escolas eficientes conforme a Tabela 7 e a Figura 30. As razões das eficiências dessas escolas foram explicadas na seção 4.2.1.

**Tabela 7: Escolas eficientes DEA 2, 5º ano, Conglomerado 1.**

Escolas eficientes	Prof. total	Total alunos	Prova Brasil
1 - EC 305 SUL	16	319	53,69
2 - EC 411 NORTE	8	194	44,51
3 - CEF 02 DE Brazlândia	53	1600	32,86
4 - EC KANEGAE	6	119	34,65
5 - EC AGUILHADA	6	128	25,76
6 - CEF 02 DA ESTRUTURAL	32	1509	30,61

**Figura 30: DEA 2, Conglomerado 1, 5º ano.**



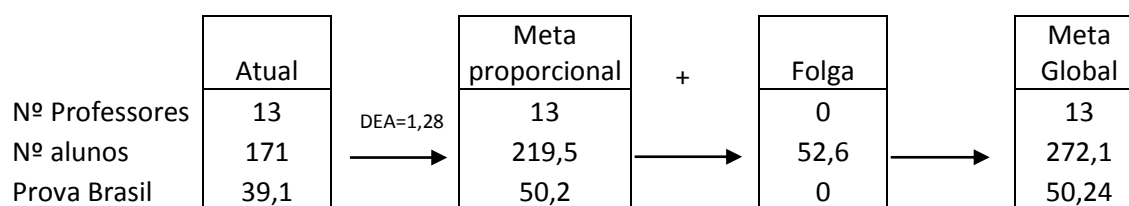
Para cada escola considerada ineficientes, a DEA identificada as escolas de referências e indica possibilidade de aumento da produtividade baseadas na projeção do plano de operação executado pela escola na fronteira de eficiência, gerando dois conjuntos de metas:

- meta proporcional, obtida pela expansão proporcional máxima da produção, representada na medida de eficiência obtida (distância à fronteira);

- meta global, que acresce à meta proporcional eventuais folgas na produção de resultados.

Para exemplificar a construção destas metas, analisou-se os resultados obtidos pela escola (ineficiente) EC 209SUL. A EC 209 Sul possui medida de eficiência de 0,78 e tem a Medida DEA igual a 1,28 indicando que pode aumentar a proporcionalmente todos os resultados em 28% sem acréscimo dos recursos. Isto gera uma possibilidade de “aumento proporcional” da produção que chamou-se “meta proporcional”. Esta projeção radial (proporcional) na direção da fronteira gerou folga na variável *número de alunos* indicando possibilidade de expansão da escola mantido o número de professores. Esta expansão, porém, depende de variáveis não consideradas neste trabalho (demanda, por exemplo). Agregando-se as folgas à meta proporcional, obtém-se a meta global que consiste no ponto da fronteira de eficiência no qual a EC209SUL foi projetada e que é definido pelas suas referências. A Figura 31 mostra os valores observados na EC 209 Sul e as metas projetadas pelo modelo DEA2.

**Figura 31: EC 209 sul ineficiente para atingir a eficiência**



A escola EC 209 Sul tem como referências as escolas EC 305 Sul e EC 411 e a sua meta global é dada por:

$$EC\ 209\ Sul_0^* = 0,625\ EC\ 305\ Sul + 0,375\ EC\ 411\ Norte$$

Esta meta global constitui uma escola virtual e a trajetória entre os valores atuais da  $EC\ 209\ Sul_0^*$  e sua meta global identificou ações que conduzem a aumento da produtividade. A Tabela 16 do Anexo apresenta as metas proporcionais e globais para todas as escolas ineficientes avaliadas.

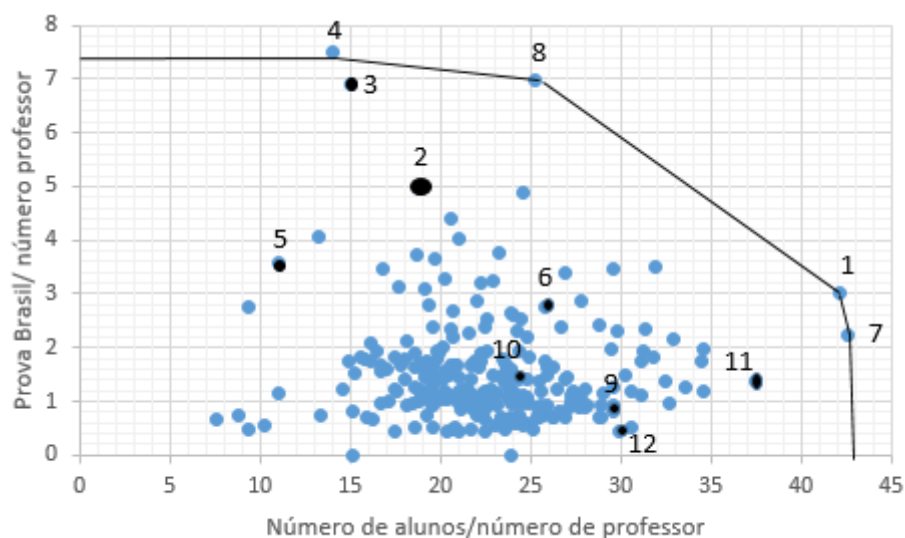
## 5.2 5º ano, conglomerado 2

A aplicação do modelo DEA 2 as escolas do 5º ano, conglomerado 2 identificou 12 escolas eficientes explicitadas na Tabela 8 e na Figura 32.

**Tabela 8: DEA 2, 5º ano, conglomerado 2**

Escolas eficientes	Prof. Total	Total alunos	Prova Brasil
1 - EC 08 DE Taguatinga	12	506	36,25
2 - CEI 05 DE Taguatinga	9	169	44,58
3 - EC SANTA HELENA	6	90	41,38
4 - EC BROCHADO DA ROCHA	4	56	29,96
5 - EC IPE	14	154	50,02
6 - EC 03 DO NUCLEO Núcleo Bandeirante	17	438	46,73
7 - EC 57 DE Ceilândia	16	682	35,51
8 - EC CORREGO DO BARREIRO	5	126	34,96
9 - CAIC Santa Maria	54	1595	44,33
10 - EC 104 DE SAO SEBASTIAO	34	825	49,84
11 - EC VILA NOVA	25	938	34,24
12 - CED DONA AMERICA GUIMARAES	78	2331	33,26

**Figura 32: DEA 2, Conglomerado 2, 5º ano**



A EC BROCHADO DA ROCHA é eficiente por ser a menor escola e maior relação entre Prova Brasil por número de professor, a CED DONA AMERICA GUIMARAES é eficiente por ser a maior escola, a EC IPE é eficiente por ter a maior nota da Prova Brasil e a EC 08 DE Taguatinga é eficiente por ter maior relação entre número de alunos por número de professor. Enquanto as outras escolas são eficientes porque é melhor que quaisquer combinações convexas entre essas escolas citadas.

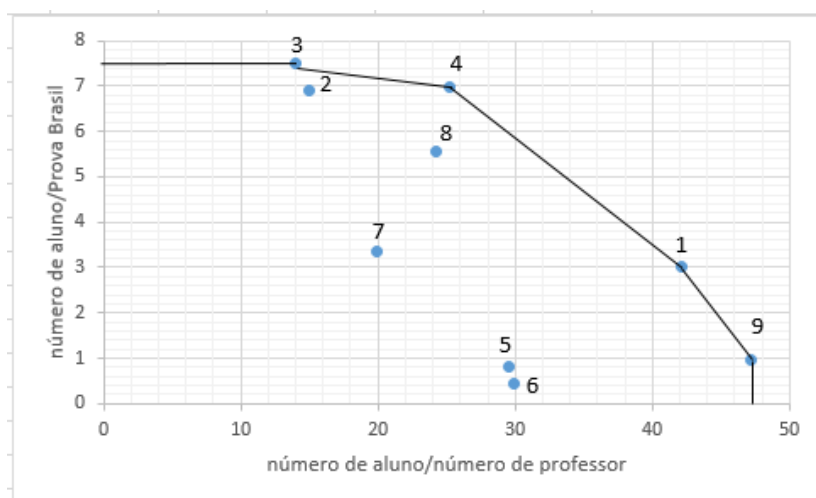
### 5.3 5º ano comparando conglomerado 1 e 2.

Para comparar as escolas dos conglomerados 1 e 2, aplicou-se novamente o modelo DEA2 ao conjunto das escolas do Ensino Fundamental I, resultando 9 escolas eficientes, três delas do Conglomerado 1 e seis do conglomerado 2. Desta forma, escolas do conglomerado 2 (menos equipadas e em regiões de renda mais baixa) mantiveram-se eficientes mesmo quando comparadas com escolas do conglomerado 1 (mais equipadas em regiões de renda mais alta). Essas escolas, mesmo em condições desfavoráveis, apresentam bom desempenho em relação ao número de alunos matriculados e seus resultados na Prova Brasil. Neste sentido, essas escolas constituem exemplos de boas práticas e as razões deste desempenho deve ser objeto de estudos futuros com um conjunto maior de variáveis. A Tabela 9 e a Figura 33 apresentam os resultados obtidos.

**Tabela 9: 5º ano escolas eficientes**

Escolas eficientes	Prof. Total	Total de alunos	Prova Brasil	Conglomerado
1 - EC 08 DE Taguatinga	12	506	36,25	2
2 - EC SANTA HELENA	6	90	41,38	2
3 - EC BROCHADO DA ROCHA	4	56	29,96	2
4 - EC CORREGO DO BARREIRO	5	126	34,96	2
5 - CAIC Santa Maria	54	1595	44,33	2
6 - CED DONA AMERICA GUIMARAES	78	2331	33,26	2
7 - EC 305 SUL	16	319	53,69	1
8 - EC 411 NORTE	8	194	44,51	1
9 - CEF 02 DA ESTRUTURAL	32	1509	30,61	1

**Figura 33: Escolas eficientes**



## 5.4 9º ano conglomerado 1.

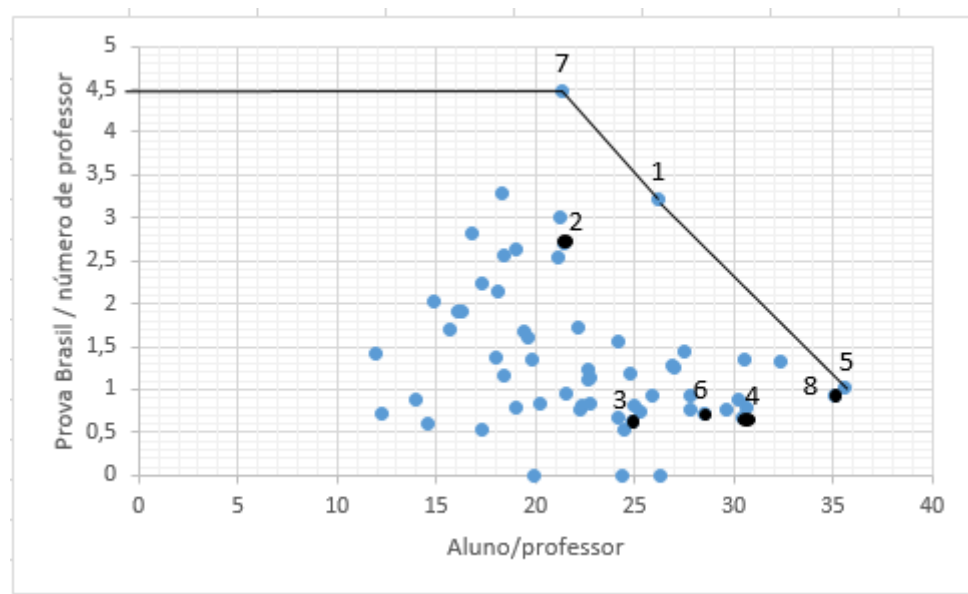
### Escolas eficientes do 9º ano, conglomerado 1

Aplicando-se o modelo DEA2 às escolas do 9º ano conglomerados 1 obteve-se 8 escolas eficientes conforme a Tabela 10 e a Figura 34.

**Tabela 10: escolas eficientes do 9º ano, conglomerado 1**

Escolas eficientes	Total prof	Total de alunos	Prova Brasil
1 - CEF 01 DE BRASILIA	17	446	54,84
2 - CEF 03 DE BRASILIA	19	407	51,55
3 - CED 07 DO Gama	82	2039	51,15
4 - CED 11 DE Ceilândia	68	2071	45,79
5 - CED 104 DO Recanto das Emas	37	1318	37,56
6 - CEF 01 DO Paranoá	68	1939	48,72
7 - CEF BONSUCESSO	15	320	67,09
8 - CED STELLA DOS CHERUBINS GUIMARAES TROIS	50	1756	46,82

**Figura 34: DEA 2, 9º ano, conglomerado 1.**



A CEF BONSUCESSO é eficiente por dominar a relação Prova Brasil por número de professor e também a ter maior nota da Prova Brasil. A CED 11 DE Ceilândia é eficiente por ser a maior escola e não poder ser descrita como combinação convexa das demais, enquanto a CED 104 do Recanto das Emas é eficiente porque domina a relação número de alunos por professor. As demais escolas são eficientes em função da adoção de retornos variáveis à escala reduzindo a região viável às combinações convexas dos planos de operação observados.

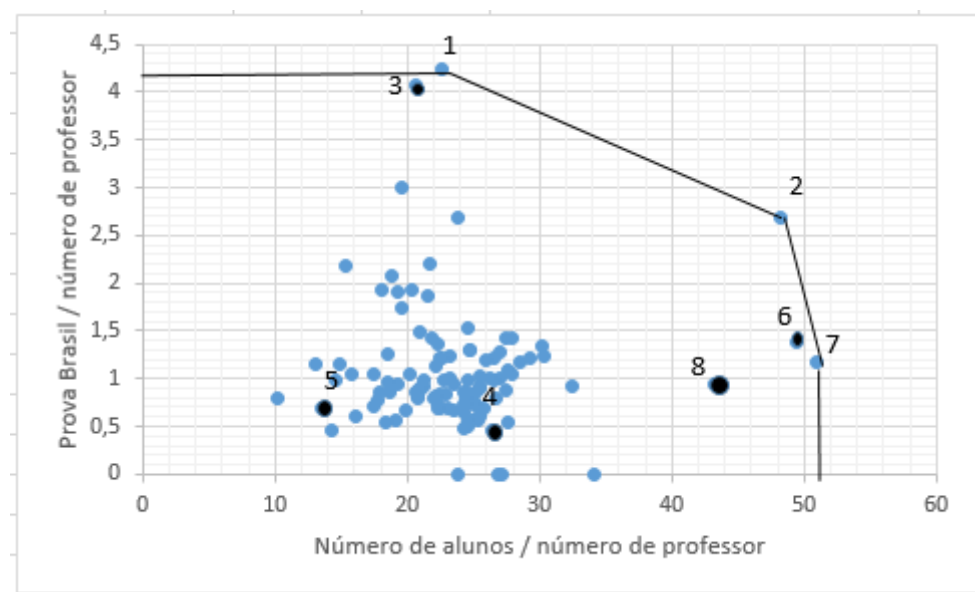
## 5.5 9º ano conglomerado 2.

Aplicando-se o modelo DEA2 às escolas do 9º ano conglomerados 2 obteve-se 8 escolas eficientes conforme a Tabela 11 e a Figura 35.

**Tabela 11: Escolas eficientes 9º ano, conglomerado 2.**

Nome	Total prof.	Total alunos	Prova Brasil
1 - CEF PONTE ALTA NORTE	13	294	55,16
2 - CEF 10 DE Taguatinga	18	869	48,37
3 - CEF JARDIM II	12	247	48,83
4 - CED 07 DE Ceilândia	93	2449	41,97
5 - CEF 301 DO Recanto das Emas	97	1315	66,04
6 - CEF 03 DO Paranoá	37	1828	51,41
7 - CEF SAO BARTOLOMEU	43	2189	50,14
8 - CED DONA AMERICA GUIMARAES	54	2331	50,48

**Figura 35: Escolas eficientes 9º ano, conglomerado 2.**



A CEF 301 DO Recanto das Emas é eficiente porque possui a maior nota na Prova Brasil, a CEF JARDIM II é eficiente por ser a menor escola, a CED 07 DE Ceilândia é eficiente por ser a maior escola, CEF SAO BARTOLOMEU é eficiente porque domina a relação número de alunos por número de professor e a CEF PONTE ALTA NORTE porquê domina a relação Prova Brasil por número de professor. Enquanto as outras escolas são eficientes pela é melhor em combinação linear entre essas escolas.

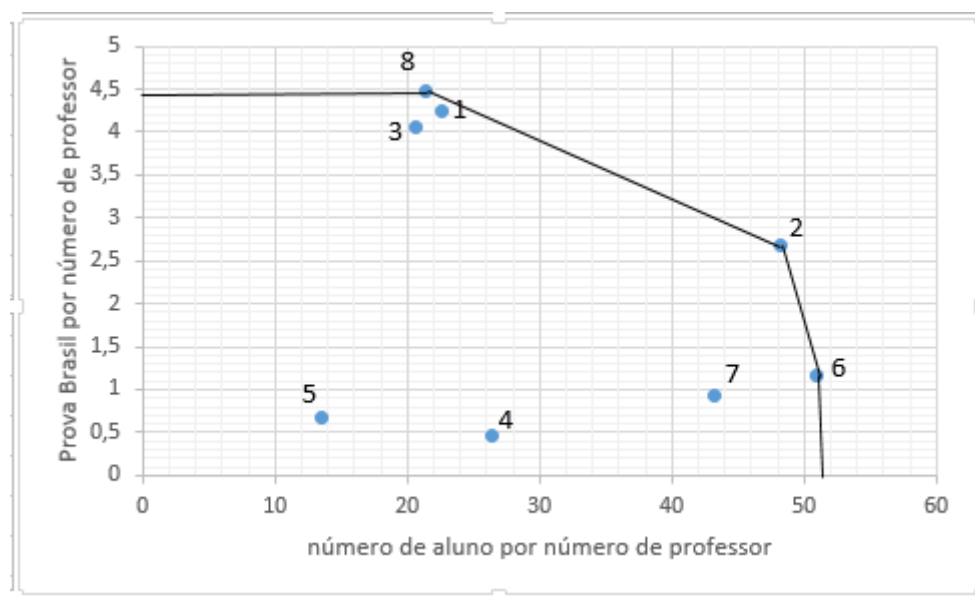
## 5.6 Comparando as escolas eficientes do conglomerado 1 e 2.

Aplicando-se o modelo DEA2 às escolas do 9º ano conglomerados 1 e 2 obteve-se 8 escolas eficientes conforme a Tabela 12 e a Figura 36.

**Tabela 12: Escolas eficientes entre conglomerado 1 e 2**

Nome	Total de Prof.	Total de alunos	Prova Brasil	Conglomerado
1 - CEF PONTE ALTA NORTE	13	294	55,16	2
2 - CEF 10 DE Taguatinga	18	869	48,37	2
3 - CEF JARDIM II	12	247	48,83	2
4 - CED 07 DE Ceilândia	93	2449	41,97	2
5 - CEF 301 DO Recanto das Emas	97	1315	66,04	2
6 - CEF SAO BARTOLOMEU	43	2189	50,14	2
7 - CED DONA AMERICA GUIMARAES	54	2331	50,48	2
8 - CEF BONSUCESSO	15	320	67,09	1

**Figura 36: Escolas eficientes do conglomerado 1 e 2.**



A CEF BONSUCESSO é eficiente por ter a maior nota da Prova Brasil e a relação Prova Brasil por número de professor, a CED 07 DE Ceilândia é eficiente por ser a maior escola, a CEF JARDIM II é eficiente por ser a menor escola e a CEF 301 DO Recanto das Emas é eficiente por ter a melhor relação número de alunos por número de professor. Enquanto as outras escolas são eficientes pela é melhor em combinação linear entre essas escolas.

## 5.7 Conclusão.

Em função das diferenças entre os grupos identificada na Análise de Conglomerados, aplicou-se a Análise Envoltória de Dados dentro de cada grupo em cada um dos segmentos (Ensino Fundamental 1 e Ensino Fundamental 2), com a mesma função de desempenho que considerou o número total de professores como insumo e o número de alunos matriculados e os resultados da Prova Brasil como produtos. Resultaram quatro aplicações de DEA que permitiram identificar escolas eficientes em cada subgrupo considerado e estabelecer ações de melhoria naquelas consideradas ineficientes.

Assim, tanto nas escolas do Ensino Fundamental 1 quanto do Ensino Fundamental 2, aplicou-se DEA em escolas com condições semelhantes de infraestrutura e renda local, permitindo maior comparabilidade e aceitação das referências.

Após analisar as escolas dentro de cada conglomerado, realizou-se uma aplicação de DEA agregando os conglomerados, tanto no Ensino Fundamental 1, quanto no Ensino Fundamental 2. Observou-se, em ambos os segmentos, que algumas escolas do conglomerado com piores condições de infraestrutura e renda se mantinham eficientes mesmo quando comparada com o universo das escolas. Essas escolas apresentam bom desempenho mesmo em condições desfavoráveis de infraestrutura e renda, qualificando-se, assim, como referências importantes do sistema de ensino, cujos processos de ensino aprendizagem devem ser estudados em função do bom desempenho tanto quantitativo quando qualitativo alcançados.



## 6 Conclusões e Recomendações

Esta pesquisa iniciou-se pela caracterização das escolas do ensino fundamental do sistema de ensino do DF. Foi observada a existência de escolas em condições diferentes tanto em relação a infraestrutura quando ao ambiente socioeconômico no qual se insere.

Em um 1º momento separou-se as escolas do Ensino Fundamental 1 (1º ao 5º ano) e do Ensino Fundamental 2 (6º ao 9º ano). Dentro de cada um destes segmentos aplicou-se Análise de Conglomerados a partir de indicadores de infraestrutura da escola e renda da região em que se insere. Destacaram-se, em ambos os segmentos dois grupos de escolas. O primeiro caracterizado por instituições melhores equipadas em regiões de renda mais alta. O segundo por escolas em piores condições de infraestrutura em regiões de renda mais baixa.

Pela análise de conglomerados tem-se que as escolas mais equipadas possuem melhores notas na Prova Brasil e, com isto, há necessidade em investir mais em instalações, televisão por sala de aula e acesso a computador. Porque quando a escola tem uma boa infraestrutura motiva os docentes e os alunos. Nota-se que nessas escolas os professores e alunos possuem condições de fazer trabalhos diferenciados. Como uma boa biblioteca esse grupo passará a ler mais tendo como consequência um rendimento melhor na Prova Brasil. Além disso, com o acesso à internet com a orientação do professor, os alunos podem consultar sites confiáveis que podem ajudá-los a ter informações precisas, como a vídeo aula de “khan academy”, “só matemática” entre outros.

Como a correlação linear entre a titulação do professor e a nota da Prova Brasil é baixa, há uma indicação que os professores com maior formação não estão influenciando na nota da Prova Brasil. Pode-se deduzir a estabilidade do professor acomoda em preparar melhor as aulas. A esse motivo, sugere-se que o professor deva ter melhores salários em função do rendimento dos alunos. Além disso, sabemos que quanto maior a formação do professor ele estará mais preparado em conteúdo. Porém, não temos os dados referentes a quantidade e a qualidade do curso do professor. Por isso, o professor deveria fazer cursos constantemente para que ele se mantenha atualizado. No caso do Distrito Federal, o professor possui a coordenação que é utilizada para preparar aula, planejar projetos da escola e corrigir as provas. Por isto, deveria ter curso sobre o conteúdo que ele irá ministrar para forçar o professor a ficar atualizado sobre o assunto. Já foi presenciado professores desatualizados e ensinando o conteúdo que não se ensina mais como a tabela de logaritmo ou ensinando errado por não estar estudando diariamente. Sem esquecer, que é o papel do docente mostrar a importância da Prova Brasil e o que significa esse resultado para os alunos. Com isso, os alunos terão mais empenho em fazer essa avaliação. E também falta o feedback do INEP para o aluno, pois os avaliados não possuem o resultado individual e isto mostra que estas pessoas que fazem a prova não enxergam esta importância.

Em função das diferenças entre os grupos identificadas na Análise de Conglomerados, aplicou-se a Análise Envoltória de Dados dentro de cada grupo em cada um dos segmentos (Ensino Fundamental 1 e Ensino Fundamental 2), com a mesma função de desempenho que considerou o número total de professores como insumo e o número de alunos matriculados e os resultados da Prova Brasil como produtos. Resultaram quatro aplicações de DEA que permitiram identificar escolas eficientes em cada subgrupo considerado e estabelecer ações de melhoria naquelas consideradas ineficientes.

Assim, tanto nas escolas do Ensino Fundamental 1 quanto do Ensino Fundamental 2, aplicou-se DEA em escolas com condições semelhantes de infraestrutura e renda local, permitindo maior comparabilidade e aceitação das referências.

Após analisar as escolas dentro de cada conglomerado, realizou-se uma aplicação de DEA agregando os conglomerados, tanto no Ensino Fundamental 1, quanto no Ensino Fundamental 2. Observou-se, em ambos os segmentos, que algumas escolas do conglomerado com piores condições de infraestrutura e renda se mantinham eficientes mesmo quando comparada com o universo das escolas. Essas escolas apresentam bom desempenho mesmo em condições desfavoráveis de infraestrutura e renda, qualificando-se, assim, como referências importantes do sistema de ensino, cujos processos de ensino aprendizagem devem ser estudados em função do bom desempenho tanto quantitativo quanto qualitativo alcançados.

A DEA sugere ações que tornam estas escolas eficientes como referência e conduzem as demais a melhorias de desempenho.

A taxa de aprovação dos estudantes não se mostrou adequada na construção dos modelos DEA, uma vez que apresenta baixa variância, não diferenciando as escolas caracterizando-se mais como uma política de ensino do que como uma medida do desempenho do aluno. Pode-se deduzir que a maioria dos sistemas escolares ainda mantém amplamente a ficção segundo a qual todas as crianças de seis anos que entram na primeira série da escola obrigatória estariam igualmente desejosas e seriam capazes de aprender a ler e a escrever em um ano. Todo mundo sabe que isso é falso, o que não impede que tal ficção permaneça no princípio da estrutura escolar, do tratamento das faixas etárias e da distribuição do programa em graus anuais. No início da escolaridade obrigatória, as diferenças de idades são as únicas que a escola aceita levar em conta. Para afrontar a formidável diversidade dos ritmos de desenvolvimento, desejou-se ignorar ou deixar por conta das dispensas por idade o fato de que, aos seis anos, certos alunos possam manifestar um nível de desenvolvimento que outros só atingirão aos sete ou oito, ao passo que outros já o haviam atingido aos quatro ou cinco. Um atraso de desenvolvimento só é considerado quando tiver originado dificuldades graves, até mesmo um fracasso. A repetição da primeira série de escolaridade obrigatória, que pretende aumentar a homogeneidade dos alunos que passam para o ano seguinte, está muito fortemente ligada à classe social, que é, assim, indiretamente considerada, por uma medida de diferenciação grosseira e cujos efeitos são duvidosos.

Apesar do pequeno número de variáveis disponíveis e da heterogeneidade das escolas, a metodologia proposta mostrou-se eficaz na descrição do sistema de ensino, na identificação dos melhores desempenhos e na construção de estratégias que conduzam as escolas ineficientes a melhorias.

## 7 Referências

- BELLONI, J. A. **Uma metodologia de avaliação da eficiência produtiva de universidades federais brasileiras**. 2000. Tese (Doutorado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2000.
- RHODES, E.L. **Data Envelopment Analysis and approaches for measuring the efficiency of decision making units with an application to program follow-through in U.S. education**. Ph.D. Dissertation, Carnegie Mellon University, 1978.
- CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, 2(6),429-444, 1978.
- BANKER, R.D.; CHARNES A.; COOPER, W.W. Some models for estimation technical and scale inefficiencies in Data Envelopment Analysis. **Management Science**, 30(9):1078-1092, 1984.
- CHARNES,A. e COOPER, W.W. Programming with linear fractional functionals. **Naval Res. Logist. Quart.**, 9, 181-185, 1962.
- FÄRE, R.; GROSSKOPF, S.; LOVELL, C.A.K. **Production frontiers**. New York, Cambridge University, 1994.
- [www.inep.gov.br](http://www.inep.gov.br)
- <http://dados.gov.br/dataset/taxas-de-rendimento-escolar-na-educacao-basica>
- MIETO NETO, F. E SANTOS, L. C. B. **Avaliação da Eficiência Produtiva dos Departamentos da UnB: Uma Aplicação de DEA**, monografia do departamento de estatística 2002.
- NORMAN, M. e STOKER, B. **Data Envelopment Analysis: The Assesment of Performance**. John Wiley & Sons, 1991.

## 8 Anexo

**Tabela 13: Nível de Desempenho dos alunos segundo a Prova Brasil**

Português

Níveis de Desempenho dos alunos em Leitura	O que os alunos conseguem fazer nesse nível
<b>Nível 0 - abaixo de 125</b>	<p>A Prova Brasil não utilizou itens que avaliam as habilidades abaixo deste nível. Os alunos localizados abaixo do nível 125 requerem atenção especial, pois, não demonstram habilidades muito elementares como as de:</p> <ul style="list-style-type: none"><li>• Localizar informação (exemplo: o personagem principal, local e tempo da narrativa);</li><li>• Identificar o efeito de sentido decorrente da utilização de recursos gráficos (exemplo: letras maiúsculas chamando a atenção em um cartaz); e</li><li>• Identificar o tema, em um texto simples e curto.</li></ul>
<b>Nível 1 - 125 a 150</b>	<p>Os alunos do 5º e 9º anos (4ª. e 8ª. séries):</p> <ul style="list-style-type: none"><li>• Localizam informações explícitas em textos narrativos curtos, informativos e anúncios;</li><li>• Identificam o tema de um texto;</li><li>• Localizam elementos como o personagem principal;</li><li>• Estabelecem relação entre partes do texto: personagem e ação; ação e tempo; ação e lugar.</li></ul>
<b>Nível 2 - 150 a 175</b>	<p>Este nível é constituído por narrativas mais complexas e incorporam outros gêneros textuais, por isto, ainda que algumas habilidades aqui apontadas já estejam listadas anteriormente, elas se mostraram mais difíceis neste intervalo.</p> <p>Além das habilidades anteriormente citadas, os alunos do 5º e 9º anos (4ª. e 8ª. séries):</p> <ul style="list-style-type: none"><li>• Localizam informação explícita. Exemplo: identificando, dentre vários personagens, o principal, e, em situações mais complexas, a partir de seleção e comparação de partes do texto;</li><li>• Identificam o tema de um texto;</li><li>• inferem informação em texto verbal (características do personagem) e não verbal (tirinha);</li><li>• interpretam pequenas matérias de jornal, trechos de enciclopédia, poemas longos e prosa poética;</li><li>• identificam o conflito gerador e finalidade do texto</li></ul>
<b>Nível 3 - 175 a 200</b>	<p>Além das habilidades anteriormente citadas, os alunos do 5º e 9º anos (4ª. e 8ª. séries):</p> <ul style="list-style-type: none"><li>• interpretam, a partir de inferência, texto não-verbal (tirinha) de maior</li></ul>

	<p>complexidade temática;</p> <ul style="list-style-type: none"> <li>identificam o tema a partir de características que tratam de sentimentos do personagem principal;</li> <li>reconhecem elementos que compõem uma narrativa com temática e vocabulário complexos.</li> </ul>
<b>Nível 4 - 200 a 225</b>	<p>Além de demonstrar todas as habilidades anteriores a partir de anedotas, fábulas e textos com linguagem gráfica pouco usual, narrativos complexos, poéticos, informativos longos ou com informação científica, os alunos do 5º e do 9º anos (4ª. E 8ª. séries):</p> <ul style="list-style-type: none"> <li>identificam, dentre os elementos da narrativa que contém discurso direto, o narrador observador;</li> <li>selecionam entre informações explícitas e implícitas as correspondentes a um personagem;</li> <li>localizam informação em texto informativo, com estrutura e vocabulário complexos;</li> <li>inferem a informação que provoca efeito de humor no texto;</li> <li>interpretam texto verbal, cujo significado é construído com o apoio de imagens, inferindo informação;</li> <li>identificam o significado de uma expressão em texto informativo;</li> <li>inferem o sentido de uma expressão metafórica e o efeito de sentido de uma onomatopeia;</li> <li>interpretam história em quadrinho a partir de inferências sobre a fala da personagem, identificando o desfecho do conflito;</li> <li>estabelecem relações entre as partes de um texto, identificando substituições pronominais que contribuem para a coesão do texto.</li> </ul>
<b>Nível 5 - 225 a 250</b>	<p>Além das habilidades anteriores, os alunos do 5º e 9º anos (4ª. e 8ª. séries):</p> <ul style="list-style-type: none"> <li>identificam o efeito de sentido decorrente do uso da pontuação (reticências);</li> <li>inferem a finalidade do texto;</li> <li>distinguem um fato da opinião relativa a este fato, numa narrativa com narrador personagem;</li> <li>distinguem o sentido metafórico do literal de uma expressão;</li> <li>reconhecem efeitos de ironia ou humor em textos variados;</li> <li>identificam a relação lógico-discursiva marcada por locução adverbial ou conjunção comparativa;</li> <li>interpretam texto com apoio de material gráfico;</li> <li>localizam a informação principal.</li> </ul> <p>Os alunos do 9º ano, neste nível, ainda:</p> <ul style="list-style-type: none"> <li>inferem o sentido de uma palavra ou expressão;</li> <li>estabelecem relação causa/consequência entre partes e elementos do texto;</li> <li>identificam o tema de textos narrativos, argumentativos e poéticos de conteúdo complexo;</li> <li>identificam a tese e os argumentos que a defendem em textos argumentativos;</li> <li>reconhecem o efeito de sentido decorrente da escolha de uma determinada palavra ou expressão.</li> </ul>
<b>Nível 6 - 250 a 275</b>	<p>Utilizando como base a variedade textual já descrita, neste nível os alunos do 5º e do 9º anos (4ª. e 8ª. séries), além de demonstrarem as habilidades anteriores:</p>

	<ul style="list-style-type: none"> <li>• localizam características do personagem em texto poético;</li> <li>• distinguem um fato da opinião relativa a este fato;</li> <li>• identificam uma definição em texto expositivo;</li> <li>• estabelecem relação causa/consequência entre partes e elementos do texto;</li> <li>• inferem a finalidade do texto a partir do suporte;</li> <li>• inferem o sentido de uma palavra ou expressão;</li> <li>• identificam a finalidade do texto;</li> <li>• identificam o assunto em um poema;</li> <li>• comparam textos que tratam do mesmo tema, reconhecendo diferentes formas de tratar a informação;</li> <li>• interpretam texto a partir de material gráfico diverso (gráficos, tabelas, etc);</li> <li>• estabelecem relações entre as partes de um texto, identificando substituições pronominais que contribuem para a coesão do texto.</li> </ul> <p>Os alunos do 9º ano (8ª. série) ainda:</p> <ul style="list-style-type: none"> <li>• estabelecem relações entre partes de um texto, reconhecendo o sentido de uma expressão que contribui para a continuidade do texto;</li> <li>• estabelecem relações lógico-discursivas presentes no texto, marcadas por conjunções, advérbios, etc;</li> <li>• reconhecem o efeito de sentido decorrente da exploração de recursos ortográficos e/ou morfosintáticos;</li> <li>• identificam o conflito gerador do enredo e os elementos que constroem a narrativa;</li> <li>• identificam a tese e o argumento que defendem em texto com a linguagem informal;</li> <li>• inferem informação a partir de um julgamento em textos narrativos longos;</li> <li>• inferem efeitos de ironia ou humor em narrativas curtas;</li> </ul> <p>inferem o sentido de uma expressão em texto narrativo longo e de vocabulário complexo</p>
<b>Nível 7 - 275 a 300</b>	<p>Além de demonstrar as habilidades dos níveis anteriores, no 5º e no 9º anos(4ª. e 8ª. séries), os alunos:</p> <ul style="list-style-type: none"> <li>• inferem informação em texto narrativo longo;</li> <li>• identificam relação lógico-discursiva marcada por locução adverbial de lugar, advérbio de tempo ou termos comparativos em textos narrativos longos, com temática e vocabulário complexos.</li> </ul> <p>Os alunos do 9º ano (8ª. série):</p> <ul style="list-style-type: none"> <li>• inferem informações implícitas em textos poéticos subjetivos, textos argumentativos com intenção irônica, fragmento de narrativa literária clássica, versão modernizada de fábula e histórias em quadrinhos;</li> <li>• reconhecem o efeito de sentido decorrente da utilização de uma</li> </ul>

	<p>determinada expressão;</p> <ul style="list-style-type: none"> <li>estabelecem relação causa/consequência entre partes e elementos do texto;</li> <li>reconhecem posições distintas entre duas ou mais opiniões relativas ao mesmo fato ou tema;</li> <li>comparam textos que tratam do mesmo tema, reconhecendo diferentes formas de tratar a informação.</li> </ul>
<b>Nível 8 -300 a 325</b>	<p>Além de demonstrar as habilidades dos níveis anteriores, no 5º e no 9º anos (4ª. e 8ª. séries), os alunos:</p> <ul style="list-style-type: none"> <li>identificam o assunto do texto em narrativas longas com vocabulário complexo;</li> <li>inferem informações em fábulas.</li> </ul> <p>Os alunos do 9º ano (8ª. série):</p> <ul style="list-style-type: none"> <li>inferem o tema de texto poético;</li> <li>inferem a finalidade de texto informativo;</li> <li>identificam a opinião do autor em texto informativo com vocabulário complexo;</li> <li>diferenciam as partes principais das secundárias de um texto;</li> <li>interpretam tabela a partir da comparação entre informações;</li> <li>inferem o sentimento do personagem em história em quadrinhos;</li> <li>estabelecem relação entre a tese e os argumentos oferecidos para sustentá-la;</li> <li>identificam a tese de um texto argumentativo;</li> <li>identificam o conflito gerador do enredo;</li> <li>reconhecem o efeito de sentido decorrente do uso da pontuação e de outras notações.</li> </ul>
<b>Nível 9 - 325 a 350</b>	<p>Além das habilidades descritas anteriormente, os alunos da 9º ano (8ª. série) localizados neste nível demonstram habilidades de leitura que envolvem compreensão global de texto; avaliação e estabelecimento de relações entre textos e partes de textos mais longos e com vocabulário complexos; inferem informações em diversos contextos; e começam a ler com compreensão textos da literatura clássica.</p>

#### Matemática

Níveis de Desempenho dos alunos em	<p>O que os alunos conseguem fazer nesse nível</p> <p>e exemplos de competência</p>
------------------------------------	---

Matemática	
Nível 0 - abaixo de 125	<p>A Prova Brasil não utilizou itens que avaliam as habilidades abaixo do nível 125.</p> <p>Os alunos localizados abaixo deste nível requerem atenção especial, pois ainda não demonstraram ter desenvolvido as habilidades mais simples apresentadas para os alunos do 5º ano como exemplo:</p> <ul style="list-style-type: none"> <li>• somar e subtrair números decimais;</li> <li>• fazer adição com reserva;</li> <li>• multiplicar e dividir com dois algarismos;</li> <li>• trabalhar com frações.</li> </ul>
Nível 1 - 125 a 150	<p>Neste nível os alunos do 5º e do 9º anos resolvem problemas de cálculo de área com base na contagem das unidades de uma malha quadriculada e, apoiados em representações gráficas, reconhecem a quarta parte de um todo.</p>
Nível 2 - 150 a 175	<p>Além das habilidades demonstradas no nível anterior, neste nível os alunos do 5º e 9º anos são capazes de:</p> <ul style="list-style-type: none"> <li>• reconhecer o valor posicional dos algarismos em números naturais;</li> <li>• ler informações e dados apresentados em gráfico de coluna;</li> <li>• interpretar mapa que representa um itinerário.</li> </ul>
<b>Nível 3 - 175 a 200</b>	<p>Além das habilidades demonstradas nos níveis anteriores, neste nível os alunos do 5º e 9º anos:</p> <ul style="list-style-type: none"> <li>• calculam resultado de uma adição com números de três algarismos, com apoio de material dourado planejado</li> <li>• localizam informação em mapas desenhados em malha quadriculada;</li> <li>• reconhecem a escrita por extenso de números naturais e a sua composição e decomposição em dezenas e unidades, considerando o seu valor posicional na base decimal;</li> <li>• resolvem problemas relacionando diferentes unidades de uma mesma medida</li> <li>• para cálculo de intervalos (dias, semanas, horas e minutos)</li> </ul>
<b>Nível 4 - 200 a 225</b>	<p>Além das habilidades descritas anteriormente, os alunos do 5º e 9º anos:</p> <ul style="list-style-type: none"> <li>• lêem informações e dados apresentados em tabela;</li> <li>• reconhecem a regra de formação de uma sequência numérica e dão continuidade a ela;</li> </ul>



	<ul style="list-style-type: none"> <li>• resolvem problemas envolvendo subtração, estabelecendo relação entre diferentes unidades monetárias;</li> <li>• resolvem situação-problema envolvendo:</li> <li>• ideia de porcentagem;</li> <li>• diferentes significados da adição e subtração;</li> <li>• adição de números racionais na forma decimal;</li> <li>• identificam propriedades comuns e diferenças entre poliedros e corpos redondos, relacionando figuras tridimensionais com suas planificações.</li> </ul>
<b>Nível 5 - 225 a 250</b>	<p>Os alunos do 5º e do 9º anos, além das habilidades já descritas:</p> <ul style="list-style-type: none"> <li>• identificam a localização/movimentação de objeto em mapas, desenhado em malha quadriculada;</li> <li>• reconhecem e utilizam as regras do sistema de numeração decimal, tais como agrupamentos e trocas na base 10 e o princípio do valor posicional;</li> <li>• calculam o resultado de uma adição por meio de uma técnica operatória;</li> <li>• lêem informações e dados apresentados em tabelas;</li> <li>• resolvem problema envolvendo o cálculo do perímetro de figuras planas, desenhadas em malhas quadriculadas;</li> <li>• resolvem problemas: utilizando a escrita decimal de cédulas e moedas do sistema monetário brasileiro;</li> <li>• estabelecendo trocas entre cédulas e moedas do sistema monetário brasileiro, em função de seus valores;</li> <li>• com números racionais expressos na forma decimal, envolvendo diferentes significados da adição ou subtração;</li> <li>• reconhecem a composição e decomposição de números naturais, na forma polinomial;</li> <li>• identificam a divisão como a operação que resolve uma dada situação-problema;</li> <li>• identificam a localização de números racionais na reta numérica.</li> </ul> <p>Os alunos do 9º ano ainda:</p> <ul style="list-style-type: none"> <li>• identificam a localização/movimentação de objeto em mapas e outras representações gráficas;</li> <li>• lêem informações e dados apresentados em gráficos de colunas;</li> <li>• conseguem localizar dados em tabelas de múltiplas entradas;</li> <li>• associam informações apresentadas em listas ou tabelas ao gráfico que as representam e vice-versa;</li> <li>• identificam propriedades comuns e diferenças entre poliedros e corpos redondos, relacionando figuras tridimensionais com suas planificações;</li> <li>• resolvem problemas envolvendo noções de porcentagem.</li> </ul>

<p><b>Nível 6 - 250 a 275</b></p>	<p>Os alunos do 5º e 9º anos:</p> <ul style="list-style-type: none"> <li>• identificam planificações de uma figura tridimensional;</li> <li>• resolvem problemas:</li> <li>• estabelecendo trocas entre cédulas e moedas do sistema monetário brasileiro, em função de seus valores;</li> <li>• envolvendo diferentes significados da adição e subtração;</li> <li>• envolvendo o cálculo de área de figura plana, desenhada em malha quadriculada;</li> <li>• reconhecem a decomposição de números naturais nas suas diversas ordens;</li> <li>• identificam a localização de números racionais representados na forma decimal na reta numérica;</li> <li>• estabelecem relação entre unidades de medida de tempo;</li> <li>• lêem tabelas comparando medidas de grandezas;</li> <li>• identificam propriedades comuns e diferenças entre figuras bidimensionais pelo número de lados e pelos tipos de ângulos;</li> <li>• reconhecem a composição e decomposição de números naturais em sua forma polinomial.</li> </ul> <p>Os alunos do 9º ano também:</p> <ul style="list-style-type: none"> <li>• reconhecem as representações decimais dos números racionais como uma extensão do sistema de numeração decimal, identificando a existência de "ordens" como décimos, centésimos e milésimos;</li> <li>• identificam a localização de números inteiros na reta numérica.</li> </ul>
<p><b>Nível 7 - 275 a 300</b></p>	<p>Os alunos do 5º e 9º anos:</p> <ul style="list-style-type: none"> <li>• resolvem problemas com números naturais envolvendo diferentes significados da multiplicação e divisão, em situação combinatória;</li> <li>• reconhecem a conservação ou modificação de medidas dos lados, do perímetro, da área em ampliação e/ou redução de figuras poligonais usando malhas quadriculadas;</li> <li>• identificam propriedades comuns e diferenças entre figuras bidimensionais pelo número de lados e tipos de ângulos;</li> <li>• identificam as posições dos lados de quadriláteros (paralelismo);</li> </ul> <p>resolvem problemas:</p> <ul style="list-style-type: none"> <li>• utilizando divisão com resto diferente de zero;</li> <li>• com apoio de recurso gráfico, envolvendo noções de porcentagem;</li> <li>• estimam medida de grandezas utilizando unidades de medida convencionais ou não;</li> <li>• estabelecem relações entre unidades de medida de tempo;</li> <li>• calculam o resultado de uma divisão por meio de uma técnica operatória;</li> </ul> <p>No 9º ano:</p> <ul style="list-style-type: none"> <li>• identificam a localização/movimentação de objeto em mapas;</li> <li>• resolvem problema com números naturais, inteiros e racionais</li> </ul>

	<p>envolvendo diferentes operações (adição, subtração, multiplicação, divisão, potenciação);</p> <ul style="list-style-type: none"> <li>• calculam o valor numérico de uma expressão algébrica, incluindo potenciação;</li> <li>• interpretam informações apresentadas por meio de coordenadas cartesianas;</li> <li>• identificam um sistema de equações do 1º grau que expressa um problema</li> </ul>
<b>Nível 8 -300 a 325</b>	<p>Os alunos do 5º e do 9º anos:</p> <ul style="list-style-type: none"> <li>• resolvem problemas;</li> <li>• envolvendo o cálculo do perímetro de figuras planas;</li> <li>• desenhadas em malhas quadriculadas;</li> <li>• envolvendo o cálculo de área de figuras planas, desenhadas em malha quadriculada;</li> <li>• utilizando porcentagem;</li> <li>• utilizando unidades de medida padronizadas como km/m/cm/mm, kg/g/mg, l/ml;</li> <li>• com números racionais expressos na forma decimal, envolvendo operações de adição e subtração;</li> <li>• estimam a medida de grandezas utilizando unidades de medida convencional ou não;</li> <li>• lêem informações e dados apresentados em gráficos de coluna;</li> <li>• identificam a localização de números racionais representados na forma decimal na reta numérica.</li> </ul>
<b>Nível 9 - 325 a 350</b>	<p>Neste nível, os alunos do 5º e 9º anos:</p> <ul style="list-style-type: none"> <li>• reconhecem a conservação ou modificação de medidas dos lados, do perímetro, da área em ampliação e/ou redução de figuras poligonais usando malhas quadriculadas;</li> <li>• identificam fração como representação que pode estar associada a diferentes significados;</li> <li>• resolvem equações do 1º grau com uma incógnita;</li> </ul> <p>identificam diferentes representações de um mesmo número racional;</p> <ul style="list-style-type: none"> <li>• calculam a área de um polígono desenhado em malha quadriculada;</li> <li>• reconhecem a representação numérica de uma fração a partir do preenchimento de partes de uma figura.</li> </ul> <p>No 9º ano os alunos também:</p> <ul style="list-style-type: none"> <li>• reconhecem círculo/circunferência, seus elementos e algumas de suas relações;</li> <li>• realizam conversão e somas de medidas de comprimento;</li> <li>• identificam a expressão algébrica que expressa uma regularidade observada em seqüências de números ou figuras;</li> <li>• resolvem problemas utilizando relações entre diferentes unidades de medida;</li> </ul>

	<ul style="list-style-type: none"> <li>• resolvem problemas que envolvam equação do 2º grau;</li> <li>• identificam fração como representação que pode estar associada a diferentes significados;</li> <li>• resolvem problemas:</li> <li>• envolvendo a escrita decimal de cédulas e moedas do sistema monetário brasileiro, utilizando várias operações (adição, subtração, multiplicação e divisão);</li> <li>• utilizando as relações métricas do triângulo retângulo;</li> <li>• reconhecem que as imagens de uma figura construída por uma transformação homotética são semelhantes, identificando propriedades e/ou medidas que se modificam ou não se alteram.</li> </ul>
<b>Nível 10 - 350 a 375</b>	<p>Além das habilidades demonstradas nos níveis anteriores, neste nível, os alunos do 5º e 9º anos:</p> <ul style="list-style-type: none"> <li>• estimam a medida de grandezas utilizando unidades de medida convencional ou não;</li> <li>• identificam propriedades comuns e diferenças entre poliedros e corpos redondos, relacionando figuras tridimensionais com suas planificações;</li> <li>• calculam o resultado de uma multiplicação ou divisão de números naturais.</li> </ul> <p>No 9º ano os alunos também:</p> <ul style="list-style-type: none"> <li>• resolvem problemas envolvendo:</li> <li>• o cálculo de área e perímetro de figuras planas;</li> <li>• o cálculo do perímetro de figuras planas, desenhadas em malha quadriculada;</li> <li>• ângulos, inclusive utilizando a Lei Angular de Tales e utilizando o Teorema de Pitágoras;</li> <li>• noções de volume;</li> <li>• relações métricas do triângulo retângulo a partir de apoio gráfico significativo;</li> <li>• reconhecem as diferentes representações de um número racional;</li> <li>• estabelecem relação entre frações próprias e impróprias, as suas representações decimais, assim como localizam-nas na reta numérica;</li> <li>• efetuam cálculos simples com valores aproximados de radicais;</li> <li>• identificam uma equação ou inequação do 1º grau que expressa um problema;</li> <li>• interpretam informações apresentadas por meio de coordenadas cartesianas;</li> <li>• reconhecem as representações dos números racionais como uma extensão do sistema de numeração decimal, identificando a existência de “ordens” como décimos, centésimos e milésimos;</li> <li>• identificam relação entre quadriláteros por meio de suas</li> </ul>

	<p>propriedades;</p> <ul style="list-style-type: none"> <li>• efetuam cálculos com números inteiros, envolvendo as operações (adição;</li> </ul> <p>subtração; multiplicação; divisão e potenciação);</p> <ul style="list-style-type: none"> <li>• identificam quadriláteros observando as posições relativas entre seus lados (paralelos, concorrentes, perpendiculares);</li> <li>• identificam frações equivalentes;</li> <li>• efetuam somatório e cálculo de raiz quadrada;</li> <li>• efetuam operações com expressões algébricas;</li> <li>• identificam as medidas que não se alteram (ângulos) e as que se modificam (perímetro, lados e área) em transformações (ampliações ou reduções) de figuras poligonais usando malhas quadriculadas;</li> <li>• reconhecem ângulos como mudança de direção ou giros,</li> <li>• identificando ângulos retos e não-retos.</li> </ul>
<b>Nível 11 - 375 a 400</b>	<p>Além das habilidades demonstradas nos níveis anteriores, neste nível os alunos do 9º ano:</p> <ul style="list-style-type: none"> <li>• reconhecem círculo/circunferência, seus elementos e algumas de suas relações;</li> <li>• identificam propriedades de triângulos pela comparação de medidas de lados e ângulos;</li> <li>• efetuam operações com números racionais, envolvendo a utilização de parênteses (adição, subtração, multiplicação, divisão, potenciação);</li> <li>• reconhecem expressão algébrica que representa uma função a partir de uma tabela;</li> <li>• reconhecem figuras semelhantes mediante o reconhecimento de relações de proporcionalidade;</li> <li>• identificam: <ul style="list-style-type: none"> <li>• a localização de números racionais na reta numérica;</li> <li>• propriedades de triângulos pela comparação de medidas de lados e ângulos;</li> <li>• propriedades comuns e diferenças entre figuras bidimensionais e tridimensionais, relacionando-as com as suas planificações;</li> <li>• a relação entre as representações algébrica e geométrica de um sistema de equações do 1º grau;</li> </ul> </li> <li>• resolvem problemas: <ul style="list-style-type: none"> <li>• envolvendo noções de volume;</li> <li>• envolvendo porcentagem;</li> <li>• utilizando propriedades dos polígonos (soma de seus ângulos internos, número de diagonais, cálculo da medida de cada ângulo interno nos polígonos regulares);</li> </ul> </li> <li>• utilizando relações métricas do triângulo retângulo;</li> </ul>

	<ul style="list-style-type: none"> <li>interpretando informações apresentadas em tabelas e/ou gráficos.</li> </ul>
<b>Nível 12 - 400 a 425</b>	<p>Além das habilidades demonstradas nos níveis anteriores, neste nível os alunos do 9º ano:</p> <ul style="list-style-type: none"> <li>identificam ângulos retos e não -retos;</li> <li>identificam a expressão algébrica que expressa uma regularidade observada em seqüências de números ou figuras (padrões);</li> <li>calculam o diâmetro de circunferências concêntricas;</li> <li>resolvem problemas:</li> <li>envolvendo equação do 2º grau;</li> <li>utilizando propriedades dos polígonos (soma de seus ângulos internos, número de diagonais, cálculo da medida de cada ângulo interno nos polígonos regulares);</li> <li>envolvendo variação proporcional, direta ou inversa, entre grandezas.</li> </ul>

Pela correlação temos:

**Tabela 14: Correlação 5º ano Conglomerado 1**

	Prova Brasil	Sem pós	Total prof.	% Sem pós	profcomPG	Total alunos	aluno por prof	Aprovação	Instalação	TV/SALA	comp	dea511	dea512
Prova Brasil	1,00	-0,32	-0,28	-0,14	0,14	-0,34	-0,22	0,40	0,09	-0,42	0,00	-0,29	0,59
Sem pós	-0,32	1,00	0,85	0,57	-0,57	0,79	0,13	-0,41	0,33	0,06	-0,05	0,19	0,00
Total prof.	-0,28	0,85	1,00	0,17	-0,17	0,86	0,07	-0,44	0,42	0,06	-0,07	0,09	0,00
% Sem pós	-0,14	0,57	0,17	1,00	-1,00	0,20	0,08	-0,19	-0,03	-0,06	0,00	0,02	-0,15
profcomPG	0,14	-0,57	-0,17	-1,00	1,00	-0,20	-0,08	0,19	0,03	0,06	0,00	-0,02	0,15
Total alunos	-0,34	0,79	0,86	0,20	-0,20	1,00	0,52	-0,32	0,34	0,13	-0,06	0,43	0,17
aluno por prof	-0,22	0,13	0,07	0,08	-0,08	0,52	1,00	0,12	-0,03	0,14	-0,01	0,80	0,35
Aprovação	0,40	-0,41	-0,44	-0,19	0,19	-0,32	0,12	1,00	0,03	-0,15	0,05	0,08	0,32
Instalação	0,09	0,33	0,42	-0,03	0,03	0,34	-0,03	0,03	1,00	-0,24	0,10	-0,02	0,18
TV/AS	-0,42	0,06	0,06	-0,06	0,06	0,13	0,14	-0,15	-0,24	1,00	-0,12	0,17	-0,22
comp	0,00	-0,05	-0,07	0,00	0,00	-0,06	-0,01	0,05	0,10	-0,12	1,00	-0,06	-0,11
dea511	-0,29	0,19	0,09	0,02	-0,02	0,43	0,80	0,08	-0,02	0,17	-0,06	1,00	0,52
dea512	0,59	0,00	0,00	-0,15	0,15	0,17	0,35	0,32	0,18	-0,22	-0,11	0,52	1,00

**Tabela 15: Escolas do 5º ano conglomerado 1 e eficientes segundo peso professor, aluno e Prova Brasil.**

Nome	Peso professor	Peso aluno	Peso Prova Brasil	Termo independente
CEF 01 DO PLANALTO	0,018212	0,000723	0,024666	1,263617
EC 102 SUL	0	0,000151	0,020545	1,151103
EC 106 NORTE	0,013327	0,000529	0,01805	0,924679
EC 108 SUL	0,02992	0,000718	0,016292	0,625096
EC 113 NORTE	0,030693	0,000737	0,016713	0,641264
EC 114 SUL	0	0,000168	0,022866	1,281139
EC 115 NORTE	0,032147	0,000772	0,017505	0,67164
EC 204 SUL	0,03554	0,000853	0,019352	0,742517
EC 206 SUL	0	0,000153	0,020825	1,166781
EC 209 SUL	0,029322	0	0,025551	0,902622
EC 302 NORTE	0,032341	0,000776	0,017611	0,675684
EC 304 NORTE	0,032147	0,000772	0,017505	0,671632
EC 304 SUL	0,029393	0	0,025613	0,904804
EC 305 SUL	0,027106	0,000651	0,01476	0,566308
CEF 306 NORTE	0	0,000143	0,019558	1,095764
EC 308 SUL	0,030343	0,000728	0,016522	0,633935
EC 312 NORTE	0,023013	0	0,020054	0,708428
EC 314 SUL	0	0,000133	0,018165	1,017731
CEF 316 NORTE	0,014144	0,000562	0,019157	0,981373
EC 316 SUL	0	0,000163	0,022259	1,247142
EC 403 NORTE	0,033806	0,000812	0,018408	0,706285
EC 405 NORTE	0,031193	0	0,027182	0,960228
EC 407 NORTE	0,026223	0	0,02285	0,807218
EC 410 SUL	0	0	0,021943	1,178058
EC 411 NORTE	0,051396	0,001124	0,01757	0,58883
EC 413 SUL	0	0	0,026646	1,430544
EC 415 NORTE	0,031687	0,000761	0,017255	0,662027
EC 416 SUL	0	0	0,026609	1,428591
EC 708 NORTE	0,037672	0,000904	0,020513	0,787063
EC DO SMU	0,023026	0	0,020065	0,708813
EC GRANJA DO TORTO	0,11443	0,002035	0,021461	0,434481
CEF PONTE ALTA DO BAIXO	0,048002	0,001152	0,026138	1,00288
EC 22 DO Gama	0	0,00017	0,02321	1,300413
CEF ENGENHO DAS LAJES	0	0,000159	0,021656	1,213321
EC 13 DE Taguatinga	0,037865	0,000909	0,020619	0,791106
EC 16 DE Taguatinga	0,014245	0,000566	0,019294	0,988384
EC 17 DE Taguatinga	0,013866	0,000551	0,018781	0,962128
EC 29 DE Taguatinga	0,034441	0,000827	0,018754	0,719556
CEF VENDINHA	0,015684	0,000623	0,021243	1,088258
CEF 02 DE Brazlândia	0,008004	0,000415	0,010234	0,683059
EC 06 DE Brazlândia	0,031948	0,000767	0,017397	0,667479



EC CHAPADINHA	0,035228	0	0,030698	1,084428
CED IRMA MARIA REGINA VELANES				
REGIS	0	0,000322	0,021415	1,462673
CEF 04 DE Sobradinho	0	0,000271	0,018	1,229448
EC 12 DE Sobradinho	0,014363	0,00057	0,019454	0,996609
EC CORREGO DE Sobradinho	0,036927	0,000887	0,020108	0,771503
EC SONHEM DE CIMA	0,031519	0	0,027466	0,970271
CEF PIPIRIPAU II	0,017168	0,000682	0,023253	1,191227
CEF RIO PRETO	0,037272	0,000895	0,020295	0,778698
CED TAQUARA	0,016042	0,000637	0,021727	1,113053
CEF 08 DE Planaltina	0,013306	0,000528	0,018022	0,923238
EC ALTA-MIR	0,04209	0,00101	0,022919	0,879374
CED OSORIO BACCHIN	0,047536	0,001141	0,025884	0,993138
EC SANTOS DUMONT	0,045673	0,001096	0,02487	0,95423
CED DO PAD-DF	0,011692	0,000464	0,015837	0,81128
EC NATUREZA	0,116422	0,00207	0,021835	0,442045
CED AGROURBANO IPE Riacho Fundo	0	0,000174	0,02371	1,328424
EC KANEGAE	0,107581	0,001339	0,024263	0,479037
CAIC JUSCELINO KUBITSCHK	0	0,000155	0,021164	1,185747
EC 04 DO NUCLEO Núcleo Bandeirante	0,033438	0,000803	0,018208	0,698598
EC 12 DE Ceilândia	0,032601	0,000783	0,017752	0,681111
EC 22 DE Ceilândia	0,045925	0,001004	0,0157	0,526148
EC 35 DE Ceilândia	0,011335	0,00045	0,015352	0,786447
EC 01 DO Guará	0,032172	0,000772	0,017518	0,672147
EC 02 DO Guará	0,035884	0,000861	0,01954	0,749701
EC 05 DO Guará	0,029489	0,000708	0,016058	0,616106
EC 501 DE Samambaia	0,014141	0,000562	0,019153	0,981154
CED CASA GRANDE	0	0,000177	0,024102	1,350348
EC 102 DO Recanto das Emas	0,036456	0,000875	0,019851	0,761662
CEF 01 DO Lago Norte	0,01244	0,000494	0,016849	0,863137
EC 01 SHI-SUL	0,031867	0,000765	0,017353	0,665786
EC JARDIM BOTANICO	0,031039	0,000745	0,016901	0,648477
CEF 06 DE BRASILIA	0,01548	0,000615	0,020966	1,074048
CEF NOVA BETANIA	0,015065	0,000598	0,020405	1,045294
CEF 403 DE Santa Maria	0,012438	0,000494	0,016846	0,862984
EC AGUILHADA	0,140703	0,002502	0,026388	0,534235
CEF 101 DO Recanto das Emas	0,013087	0,00052	0,017725	0,908045
EC VILA DO BOA	0,091783	0,001632	0,017214	0,348491
CEF BONSUCESSO	0,037668	0,000904	0,020511	0,786986
EC ASPALHA	0,036143	0,000868	0,019681	0,755111
EC VALE DO SOL	0,041921	0,001006	0,022827	0,875842
EC BASEVI	0	0,00019	0,025935	1,453086
EC CORREGO DO ARROZAL	0,03205	0	0,027929	0,986605
EC 01 DE ITAPOA	0,012479	0,000496	0,016901	0,865827
CEF 02 DA ESTRUTURAL	0,03125	0,000663	0	0
EC 08 DE Taguatinga	0,083333	0,001685	0,004071	0

EC SANTA HELENA	0,137927	0	0,024167	0,172435
EC BROCHADO DA ROCHA	1,25	0,017857	0	-4
EC CORREGO DO BARREIRO	0,113676	0,002021	0,02132	0,431619
CAIC Santa Maria	0,008105	0,000322	0,010977	0,562346
CED DONA AMERICA GUIMARAES	0,007666	0,000429	0	0,402048

---

**Tabela 16: Escolas eficientes segundo escolas 5º ano conglomerado 1.**

Nome	EC 305 SUL	EC 411 NORTE	CEF 02 DA ESTRUTURAL	EC 08 DE Taguatinga	EC SANTA HELENA	EC BROCHADO DA ROCHA	EC CORREGO DO BARREIRO	CAIC Santa Maria	CED DONA AMERICA GUIMARAES
CEF 01 DO PLANALTO	0,573751	0	0,099885	0	0	0	0	0,326364	0
EC 102 SUL	0,980267	0	0	0	0	0	0	0,019733	0
EC 106 NORTE	0,947403	0	0,045394	0	0	0	0	0,007202	0
EC 108 SUL	0,821881	0,160413	0,017706	0	0	0	0	0	0
EC 113 NORTE	0,843563	0,104291	0,052146	0	0	0	0	0	0
EC 114 SUL	0,941764	0	0	0	0	0	0	0,058237	0
EC 115 NORTE	0,87472	0,125187	9,33E-05	0	0	0	0	0	0
EC 204 SUL	0,08878	0,899147	0,012073	0	0	0	0	0	0
EC 206 SUL	0,969278	0	0	0	0	0	0	0,030722	0
EC 209 SUL	0,625	0,375	0	0	0	0	0	0	0
EC 302 NORTE	0,189124	0,748917	0,061959	0	0	0	0	0	0
EC 304 NORTE	0,734798	0,218468	0,046734	0	0	0	0	0	0
EC 304 SUL	0,125	0,875	0	0	0	0	0	0	0
EC 305 SUL	1	0	0	0	0	0	0	0	0
CEF 306 NORTE	0,894478	0	0	0	0	0	0	0,105522	0
EC 308 SUL	0,581062	0,404292	0,014646	0	0	0	0	0	0
EC 312 NORTE	0,875	0,125	0	0	0	0	0	0	0
EC 314 SUL	0,969247	0	0	0	0	0	0	0,030753	0
CEF 316 NORTE	0,870701	0	0,041516	0	0	0	0	0,087783	0
EC 316 SUL	0,946033	0	0	0	0	0	0	0,053967	0
EC 403 NORTE	0,719681	0,270213	0,010106	0	0	0	0	0	0
EC 405 NORTE	0,875	0,125	0	0	0	0	0	0	0
EC 407 NORTE	0,625	0,375	0	0	0	0	0	0	0

EC 410 SUL	1	0	0	0	0	0	0	0	0
EC 411 NORTE	0	1	0	0	0	0	0	0	0
EC 413 SUL	1	0	0	0	0	0	0	0	0
EC 415 NORTE	0,734648	0,176901	0,088451	0	0	0	0	0	0
EC 416 SUL	1	0	0	0	0	0	0	0	0
EC 708 NORTE	0,984596	0,01027	0,005135	0	0	0	0	0	0
EC DO SMU	0,75	0,25	0	0	0	0	0	0	0
EC GRANJA DO TORTO	0	0,768205	0	0,099341	0	0	0,132454	0	0
CEF PONTE ALTA DO BAIXO	0,469481	0,35368	0,17684	0	0	0	0	0	0
EC 22 DO Gama	0,700688	0	0	0	0	0	0	0,299312	0
CEF ENGENHO DAS LAJES	0,454115	0	0	0	0	0	0	0,545885	0
EC 13 DE Taguatinga	0,173824	0,71745	0,108725	0	0	0	0	0	0
EC 16 DE Taguatinga	0,841682	0	0,09164	0	0	0	0	0,066678	0
EC 17 DE Taguatinga	0,92446	0	0,039569	0	0	0	0	0,035971	0
EC 29 DE Taguatinga	0,341337	0,522442	0,136221	0	0	0	0	0	0
CEF VENDINHA	0,587561	0	0,16694	0	0	0	0	0,245499	0
CEF 02 DE Brazlândia	0	0	0,352195	0	0	0	0	0,366626	0,281179
EC 06 DE Brazlândia	0,532736	0,144842	0,322421	0	0	0	0	0	0
EC CHAPADINHA	0,375	0,625	0	0	0	0	0	0	0
CED IRMA MARIA REGINA VELANES									
REGIS	0	0	0	0	0	0	0	0,903554	0,096446
CEF 04 DE Sobradinho	0	0	0	0	0	0	0	0,716577	0,283423
EC 12 DE Sobradinho	0,783235	0	0,192594	0	0	0	0	0,024171	0
EC CORREGO DE Sobradinho	0,379066	0,538956	0,081978	0	0	0	0	0	0
EC SONHEM DE CIMA	0,25	0,75	0	0	0	0	0	0	0
CEF PIPIRIPAU II	0,795917	0	0,170688	0	0	0	0	0,033394	0
CEF RIO PRETO	0,903839	0,02244	0,07372	0	0	0	0	0	0
CED TAQUARA	0,33085	0	0,473986	0	0	0	0	0,195164	0

CEF 08 DE Planaltina	0,423575	0	0,495643	0	0	0	0	0,080782	0
EC ALTA-MIR	0,165393	0,723072	0,111536	0	0	0	0	0	0
CED OSORIO BACCHIN	0,351226	0,474183	0,174591	0	0	0	0	0	0
EC SANTOS DUMONT	0,131311	0,829126	0,039563	0	0	0	0	0	0
CED DO PAD-DF	0,039521	0	0,386281	0	0	0	0	0,574197	0
EC NATUREZA	0	0,829917	0	0,072893	0	0	0,09719	0	0
CED AGROURBANO IPE Riacho Fundo	0,751321	0	0	0	0	0	0	0,248679	0
EC KANEGAE	0	0,248967	0	0	0,253098	0	0,497934	0	0
CAIC JUSCELINO KUBITSCHK	0,536322	0	0	0	0	0	0	0,463678	0
EC 04 DO NUCLEO Núcleo Bandeirante	0,231044	0,595971	0,172985	0	0	0	0	0	0
EC 12 DE Ceilândia	0,308429	0,419381	0,27219	0	0	0	0	0	0
EC 22 DE Ceilândia	0	0,660709	0,332142	0,007149	0	0	0	0	0
EC 35 DE Ceilândia	0,447594	0	0,226884	0	0	0	0	0,325523	0
EC 01 DO Guarά	0,678668	0,255888	0,065444	0	0	0	0	0	0
EC 02 DO Guarά	0,596	0,227667	0,176333	0	0	0	0	0	0
EC 05 DO Guarά	0,275075	0,483283	0,241642	0	0	0	0	0	0
EC 501 DE Samambaia	0,522054	0	0,09827	0	0	0	0	0,379676	0
CED CASA GRANDE	0,606574	0	0	0	0	0	0	0,393426	0
EC 102 DO Recanto das Emas	0,296821	0,468786	0,234393	0	0	0	0	0	0
CEF 01 DO Lago Norte	0,223349	0	0,114216	0	0	0	0	0,662436	0
EC 01 SHI-SUL	0,408429	0,39438	0,19719	0	0	0	0	0	0
EC JARDIM BOTANICO	0,718871	0,104086	0,177043	0	0	0	0	0	0
CEF 06 DE BRASILIA	0,546376	0	0,238078	0	0	0	0	0,215546	0
CEF NOVA BETANIA	0,449694	0	0,314165	0	0	0	0	0,236141	0
CEF 403 DE Santa Maria	0,288314	0	0,229276	0	0	0	0	0,48241	0
EC AGUILHADA	0	0,040523	0	0,12549	0	0	0,833987	0	0
CEF 101 DO Recanto das Emas	0,225607	0	0,473952	0	0	0	0	0,300441	0
EC VILA DO BOA	0	0,251586	0	0,606463	0	0	0,141951	0	0

CEF BONSUCESSO	0,462442	0,400039	0,137519	0	0	0	0	0	0
EC ASPALHA	0,094058	0,895628	0,010314	0	0	0	0	0	0
EC VALE DO SOL	0,051323	0,882452	0,066226	0	0	0	0	0	0
EC BASEVI	0,98808	0	0	0	0	0	0	0,01192	0
EC CORREGO DO ARROZAL	0,5	0,5	0	0	0	0	0	0	0
EC 01 DE ITAPOA	0,067586	0	0,065079	0	0	0	0	0,867335	0
CEF 02 DA ESTRUTURAL	0	0	1	0	0	0	0	0	0
EC 08 DE Taguatinga	0	0	0	1	0	0	0	0	0
EC SANTA HELENA	0	0	0	0	1	0	0	0	0
EC BROCHADO DA ROCHA	0	0	0	0	0	1	0	0	0
EC CORREGO DO BARREIRO	0	0	0	0	0	0	1	0	0
CAIC Santa Maria	0	0	0	0	0	0	0	1	0
CED DONA AMERICA GUIMARAES	0	0	0	0	0	0	0	0	1

---